

日本国特許
PATENT OFFICE
JAPANESE GOVERNMENT

REC'D 23 FEB 2001
WIPO
PCT
21.12.00

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application:

1999年12月21日

出願番号
Application Number:

平成11年特許願第363058号

出願人
Applicant(s):

松下電器産業株式会社

Best Available Copy

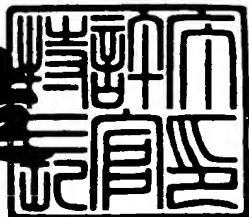
PRIORITY
DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

2001年 2月 9日

特許庁長官
Commissioner,
Patent Office

及川耕造



出証番号 出証特2001-3004716

【書類名】 特許願
【整理番号】 2030714041
【提出日】 平成11年12月21日
【あて先】 特許庁長官殿
【国際特許分類】 G06F 17/30
【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地
松下電器産業株式会社内

【氏名】 菅野 祐司

【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

【代理人】

【識別番号】 100082692

【弁理士】

【氏名又は名称】 蔵合 正博

【電話番号】 03(3519)2611

【手数料の表示】

【予納台帳番号】 013549

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9004843

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 ベクトル索引作成方法と類似ベクトル検索方法およびその装置

【特許請求の範囲】

【請求項1】 少なくともN次元実ベクトルと、その識別番号とを含む組をベクトルデータとして有限個登録したベクトルデータベースに対して、機械検索可能な索引を作成する方法であって、

ベクトル索引作成の第1ステップでは、前記ベクトルデータベース中の各ベクトルデータのN次元実ベクトルVに対して、N個の成分を予め定めた方法でm組に分けて、m個の部分ベクトル $v_1 \sim v_m$ を作成した後、部分ベクトル v_k ($k = 1 \sim m$) のノルムの分布を集計して、予め定められたD種のノルム区分のノルムの範囲を定めたノルム区分表を作成するとともに、予め定めたD個の領域中心ベクトル $p_1 \sim p_D$ に従って、前記部分ベクトル v_k の所属する領域番号dを算定し、前記部分ベクトル v_k と領域中心ベクトル p_d とのなす角の余弦($v_k \cdot p_d$) / ($|v_k| * |p_d|$) の分布を偏角分布として集計して、予め定めたC種の偏角区分の偏角の範囲を記録した偏角区分表を作成し、

ベクトル索引作成の第2ステップでは、前記ベクトルデータベース中の各ベクトルデータのN次元実ベクトルVに対して、前記第1のステップと同一の方法でN個の成分をm組に分けて、m個の部分ベクトル $v_1 \sim v_m$ を作成した後、部分空間番号bに対する部分ベクトル v_b ($b = 1 \sim m$) に対して、前記ノルム区分表を参照して前記各部分ベクトル v_b のノルムが所属するノルム区分の番号rを算定するとともに、前記第1のステップと同一の方法で、予め定めたD個の領域中心ベクトル $p_1 \sim p_D$ に従って、前記部分ベクトル v_b の所属する領域番号dを算定し、前記部分ベクトル v_b と前記領域番号dの領域の中心方向を表す領域中心ベクトル p_d とのなす角の余弦である偏角($v_b \cdot p_d$) / ($|v_b| * |p_d|$) を算定して、前記偏角区分表を参照して、所属する偏角区分の番号cを算出し、前記部分空間番号bと、前記領域番号dと、前記偏角区分番号cと、前記ノルム区分番号rと、前記部分ベクトル v_b の成分と、識別番号iとから、ベクトル索引に登録する索引登録データを算定し、

ベクトル索引作成の第3ステップでは、前記ノルム区分表と、前記偏角区分表

と、前記索引登録データとから、各部分ベクトルの識別番号と成分とを、部分空間番号 b、領域番号 d、角区分番号 c、ノルム区分番号範囲 $[r_1, r_2]$ の組を鍵にして、索可能で、かつ各ベクトルデータのベクトルの成分を、その識別番号で検索可能なベクトル索引として構成することを特徴とするベクトル索引作成方法。

【請求項 2】 少なくとも N 次元実ベクトルと、その識別番号とを含む組をベクトルデータとして有限個登録したベクトルデータベースに対して、機械検索可能な索引を作成する方法であって、

ベクトル索引作成の第 1 ステップでは、前記ベクトルデータベース中の各ベクトルデータの N 次元実ベクトル V に対して、N 個の成分を予め定めた方法で m 組に分けて、m 個の部分ベクトル $v_1 \sim v_m$ を作成した後、各部分空間番号 b に対する部分ベクトル v_b ($b = 1 \sim m$) のノルムの分布を集計して、予め定められた D 種のノルム区分のノルムの範囲を定めたノルム区分表を作成するとともに、予め定めた D 個の領域中心ベクトル $p_1 \sim p_D$ に従って、前記部分ベクトル v_b の所属する領域番号 d を算定し、前記部分ベクトル v_b と領域中心ベクトル p_d とのなす角の余弦 $(v_b \cdot p_d) / (|v_b| * |p_d|)$ の分布を偏角分布として集計して、予め定めた C 種の偏角区分の偏角の範囲を記録した偏角区分表を作成し、

ベクトル索引作成の第 2 ステップでは、前記ベクトルデータベース中の各ベクトルデータの N 次元実ベクトル V に対して、前記第 1 ステップと同一の方法で N 個の成分を m 組に分けて、m 個の部分ベクトル $v_1 \sim v_m$ を作成した後、前記部分空間 b に対する部分ベクトル v_b ($b = 1 \sim m$) に対して、前記ノルム区分表を参照して前記各部分ベクトル v_b のノルムが所属するノルム区分の番号 r を算定するとともに、前記第 1 ステップと同一の方法で、予め定めた D 個の領域中心ベクトル $p_1 \sim p_D$ に従って、前記部分ベクトル v_b の所属する領域番号 d を算定し、前記部分ベクトル v_b と前記領域番号 d の領域の中心方向を表す領域中心ベクトル p_d とのなす角の余弦である偏角 $(v_b \cdot p_d) / (|v_b| * |p_d|)$ を算定して、前記偏角区分表を参照して、所属する偏角区分の番号 c を算出し、前記算定したノルム区分番号 r に対応するノルム区分のノルムの最高値から、前記算定した部分ベクトル v_b の各成分 v_{bj} に対して、 v_{bj} が所属する予め定めた範囲

の成分区分番号 w_j を算定し、前記部分空間番号 b と、前記領域番号 d と、前記偏角区分番号 c と、前記ノルム区分番号 r と、前記成分区分番号 w_j の列と、識別番号 i とから、ベクトル索引に登録する索引登録データを算定し、

ベクトル索引作成の第3ステップでは、前記ノルム区分表と、前記偏角区分表と、前記索引登録データとから、各部分ベクトルの識別番号と成分とを、部分空間番号 b 、領域番号 d 、偏角区分番号 c 、ノルム区分番号範囲 $[r_1, r_2]$ の組を鍵にして、検索可能で、かつ各ベクトルデータのベクトルの成分を、その識別番号で検索可能なベクトル索引として構成することを特徴とするベクトル索引作成方法。

【請求項3】 前記ベクトル索引作成の第1および第2のステップにおいて、
 N 次元ベクトル V の全成分を取り出すように、 V の先頭の成分から順に、 N/m 個、あるいは $(N/m) + 1$ 個の成分を取り出して、部分ベクトルを作成する請求項1または2記載のベクトル索引作成方法。

【請求項4】 前記ベクトル索引作成の第1のステップにおいて、ノルム区分表を作成する際に、ノルム分布の集計結果をもとに、各ノルム区分に相当するノルム範囲に所属する部分ベクトルの数が、できる限り均一になるようにノルム区分を定める請求項1記載のベクトル索引作成方法。

【請求項5】 前記ベクトル索引作成の第1のステップにおいて、偏角区分表を作成する際に、偏角分布の集計結果をもとに、各偏角区分に相当する偏角範囲に所属する部分ベクトルの数が、できる限り均一になるように偏角区分を定める請求項1記載のベクトル索引作成方法。

【請求項6】 前記ベクトル索引作成の第1および第2のステップにおいて、部分ベクトル v_b の領域番号を、予め定めた D 個の領域中心ベクトル $p_1 \sim p_D$ のうち、 p_d と v_b のなす角の余弦 $(v_b \cdot p_d) / (|v_b| * |p_d|)$ がもっとも大きくなるような領域中心ベクトル p_d の番号 d として求める請求項1または2記載のベクトル索引作成方法。

【請求項7】 前記ベクトル索引作成の第3のステップにおいて、部分空間番号 b 、領域番号 d 、偏角区分番号 c 、ノルム区分番号 r を結合した番号 $(b * N_d * N_c * N_r) + (d * N_c * N_r) + (c * N_r) + r$ を鍵とし、ベクトルの識別

番号 i と成分とを検索可能な探索木と、ベクトルデータの識別番号を添字とする、各部分ベクトルの前記探索木の鍵を記録した表とを作成し、ベクトル索引の一部とする請求項1から6のいずれかに記載のベクトル索引作成方法。

【請求項8】 前記ベクトル索引作成の第2のステップにおいて、成分が $\{-1, 0, +1\}$ のうちのいずれかである、0ベクトルでない全てのベクトル $(0, \dots, 0, +1) \sim (-1, \dots, -1)$ を正規化したベクトルを領域中心ベクトルとして用いる請求項1から7のいずれかに記載のベクトル索引作成方法。

【請求項9】 検索条件としてN次元実ベクトルの質問ベクトルQと、内積下限値 α 、および最大取得ベクトル数Lとを指定して、少なくともN次元実ベクトルとそのID番号との組を有限個登録したベクトルデータから作成したベクトル索引を検索して、前記ベクトルデータベースのベクトルデータ (i, V) で、前記質問ベクトルQとの内積の値 $V \cdot Q$ が前記内積下限値 α より大きいものについて、その識別番号 i と、QとVの内積の組 $(i, V \cdot Q)$ を最大L個求める類似ベクトルの検索方法であって、

類似ベクトル検索の第1ステップでは、前記質問ベクトルQに対して、QのN個の成分を、前記ベクトル索引の作成時に用いた方法と同一の、予め定めた方法でm組に分けて、m個の部分質問ベクトル $q_1 \sim q_m$ を作成するとともに、各部分質問ベクトル q_b と対応する部分ベクトルとの内積（この内積をこれ以降「部分内積」と呼ぶ）の下限値である部分内積下限値 f_b を、指定された内積下限値 α から算定し、各部分質問ベクトル q_b ($b = 1 \sim m$) と各領域 b に対して、領域中心ベクトル p_d と前記部分質問ベクトル q_b との内積 $p_d \cdot q_b$ の値と、前記部分内積下限値 f_b と、前記ベクトル索引中のノルム区分表および偏角区分表とから、部分空間番号 b、領域番号 d において検索すべき偏角区分番号 c、ノルム区分の範囲 $[r_1, r_2]$ の組 $(c, [r_1, r_2])$ を算定し、前記算定した $(c, [r_1, r_2])$ を元に、 $(b, d, c, [r_1, r_2])$ を検索条件として前記ベクトル索引を範囲検索し、索引検索結果として、条件を満たす部分ベクトル v_b の識別番号 i と成分とを求め、前記 v_b と q_b との部分内積 $v_b \cdot q_b$ と前記部分内積下限値 f_b との差である、部分内積差分 $(v_b \cdot q_b) - f_b$ を算出して、内積差分表の識別番号 i の内積差分上限値 $S[i]$ として累算（加算）し、

類似ベクトル検索の第2ステップでは、前記内積差分表S[i]中で値の大きいものから順に、前記ベクトル索引を識別番号iで検索してベクトルデータの成分Vを求め、Vと前記質問ベクトルQとの内積V・Qから α を差し引いた内積差分値 $t = V \cdot Q - \alpha$ を算出していき、前記内積差分表中で、内積差分値を算出していない要素の最大値より大きい内積差分値を持つベクトルデータがL個以上になった時点、あるいは内積差分上限値が正である全てのベクトルデータの内積を算出した時点で、内積差分値の大きな最大L個のベクトルデータに対し、少なくとも識別番号iと内積 $t + \alpha$ との組を、検索結果として出力することを特徴とする類似ベクトル検索方法。

【請求項10】 検索条件としてN次元実ベクトルの質問ベクトルQと、距離上限値 α 、および最大取得ベクトル数Lとを指定して、少なくともN次元実ベクトルとその識別番号との組を有限個登録したベクトルデータから作成したベクトル索引を検索して、前記質問ベクトルQとの内積の値が前記距離上限値 α 以下であるような、前記ベクトルデータ中のN次元実ベクトルVの識別番号iと、QとVの距離pの組(i, p)を最大L個求める類似ベクトルの検索方法であって、

類似ベクトル検索の第1ステップでは、前記質問ベクトルQに対して、QのN個の成分を、前記ベクトル索引の作成時に用いた方法と同一の、予め定めた方法でm組に分けて、m個の部分質問ベクトル $q_1 \sim q_m$ を作成するとともに、各部分質問ベクトル q_b と対応する部分ベクトル v_b との二乗距離 $|v_b - q_b|^2$ (すなわちユークリッド距離の二乗、この値をこれ以後「部分二乗距離」と呼ぶ)の上限値である部分二乗距離上限値 f_b を、指定された距離上限値 α から算定し、各部分質問ベクトル q_b ($b = 1 \sim m$)に対して、前記部分質問ベクトル q_b と、前記部分二乗距離上限値 f_b と、前記ベクトル索引中のノルム区分表および偏角区分表とから、検索すべき部分空間番号b、領域番号d、偏角区分番号c、ノルム区分の範囲 $[r_1, r_2]$ の組(b, d, c, $[r_1, r_2]$)を系統的に生成し、前記生成した(b, d, c, $[r_1, r_2]$)を検索条件として、前記ベクトル索引を範囲検索して、索引検索結果として、条件を満たす部分ベクトル v_b の識別番号iと成分とを求め、前記部分二乗距離上限値 f_b と、 v_b と q_b との部分二乗距離 $|v_b - q_b|^2$ との差である、部分二乗距離差分 $f_b - |v_b - q_b|^2$ を算出

して、二乗距離差分表の識別番号 i の二乗距離差分上限値 $S[i]$ として累算（加算）し、

類似ベクトル検索の第2ステップでは、前記二乗距離差分表 $S[i]$ 中で値の大きいものから順に、前記ベクトル索引を識別番号 i で検索してベクトルデータの成分 V を求め、距離上限値の二乗 α^2 から V と前記質問ベクトル Q との二乗距離 $|V - Q|^2$ を差し引いた二乗距離差分値 $\alpha^2 - |V - Q|^2$ を算出していき、前記二乗距離差分表中で、二乗距離差分値を算出していない要素の最大値より大きい二乗距離差分値を持つベクトルデータが L 個以上になった時点、あるいは二乗距離差分上限値が正である全てのベクトルデータの二乗距離差分値を算出した時点で、二乗距離差分値 t の大きな最大 L 個のベクトルデータに対し、少なくとも識別番号 i と、距離 $(\alpha^2 - t)^{1/2}$ との組を、検索結果として出力することを特徴とする類似ベクトル検索方法。

【請求項11】 前記類似ベクトル検索の第1のステップにおいて、 N 次元ベクトル V の全成分を取り出すように、 V の先頭の成分から順に、 N/m 個、あるいは $(N/m) + 1$ 個の成分を取り出して、部分質問ベクトルを作成する請求項9または10記載の類似ベクトル検索方法。

【請求項12】 前記類似ベクトル検索の第1のステップにおいて、前記部分質問ベクトル q_b と対応する部分ベクトル v_b との内積の下限値である部分内積下限値 f_b を、指定された内積下限値 α から、 $f_b = \alpha |q_b|^2 / \sum (|q_b|^2)$ によって算定する請求項9記載の類似ベクトル検索方法。

【請求項13】 前記類似ベクトル検索の第1のステップにおいて、前記部分質問ベクトル q_b と対応する部分ベクトル v_b との二乗距離の上限値である部分二乗距離上限値 f_b を、指定された距離下上限値 α から、 $f_b = \alpha^2 |q_b|^2 / \sum (|q_b|^2)$ によって算定する請求項10記載の類似ベクトル検索方法。

【請求項14】 少なくとも N 次元実ベクトルと、その識別番号とを含む組をベクトルデータとして有限個登録したベクトルデータベースに対して、機械検索可能な索引を作成する装置であって、

前記ベクトルデータベース中の各ベクトルデータの N 次元実ベクトル V に対して、 N 個の成分を予め定めた方法で m 組に分けて、 m 個の部分ベクトル $v_1 \sim v_m$

を作成する部分ベクトル算定手段と、

前記作成されたm個の部分ベクトル $v_1 \sim v_m$ のうち、部分ベクトル v_k ($k = 1 \sim m$) のノルムの分布を集計して、予め定められたD種のノルム区分のノルムの範囲を定めたノルム区分表を作成するノルム分布集計手段と、

予め定めたD個の領域中心ベクトル $p_1 \sim p_D$ に従って、前記部分ベクトル v_k の所属する領域番号dを算定する領域番号算定手段と、

前記部分ベクトル v_k と領域中心ベクトル p_d とのなす角の余弦($v_k \cdot p_d$) / ($|v_k| * |p_d|$) の分布を偏角分布として集計して、予め定めたC種の偏角区分の偏角の範囲を記録した偏角区分表を作成する偏角分布集計手段と、

前記部分ベクトル算定手段が作成したm個の部分ベクトル $v_1 \sim v_m$ のうち、部分空間番号bに対する部分ベクトル v_b ($b = 1 \sim m$) に対して、前記ノルム区分表を参照して前記各部分ベクトル v_b のノルムが所属するノルム区分の番号rを算定するノルム区分番号算定手段と、

前記部分ベクトル v_b と前記領域番号算定手段が算定した領域番号dの領域の中心方向を表す領域中心ベクトル p_d とのなす角の余弦である偏角($v_b \cdot p_d$) / ($|v_b| * |p_d|$) を算定して、前記偏角区分表を参照して、所属する偏角区分の番号cを算出する偏角区分番号算定手段と、

前記部分空間番号bと、前記領域番号dと、前記偏角区分番号cと、前記ノルム区分番号rと、前記部分ベクトル v_b の成分と、識別番号iとから、ベクトル索引に登録する索引登録データを算定する索引データ算定手段と、

前記ノルム区分表と、前記偏角区分表と、前記索引登録データとから、各部分ベクトルの識別番号と成分とを、部分空間番号b、領域番号d、偏角区分番号c、ノルム区分番号範囲 $[r_1, r_2]$ の組を鍵にして、検索可能で、かつ各ベクトルデータのベクトルの成分を、その識別番号で検索可能なベクトル索引として構成する索引構成手段とを備えたことを特徴とするベクトル索引作成装置。

【請求項15】 少なくともN次元実ベクトルと、その識別番号とを含む組をベクトルデータとして有限個登録したベクトルデータベースに対して、機械検索可能な索引を作成する装置であって、

前記ベクトルデータベース中の各ベクトルデータのN次元実ベクトルVに対し

て、N個の成分を予め定めた方法でm組に分けて、m個の部分ベクトル $v_1 \sim v_m$ を作成する部分ベクトル算定手段と、

前記作成されたm個の部分ベクトル $v_1 \sim v_m$ のうち、部分空間番号bに対する部分ベクトル v_b ($b = 1 \sim m$)のノルムの分布を集計して、予め定められたD種のノルム区分のノルムの範囲を定めたノルム区分表を作成するノルム分布集計手段と、

予め定めたD個の領域中心ベクトル $p_1 \sim p_D$ に従って、前記部分ベクトル v_b の所属する領域番号dを算定する領域番号算定手段と、

前記部分ベクトル v_b と領域中心ベクトル p_d とのなす角の余弦($v_b \cdot p_d$) / ($|v_b| * |p_d|$)の分布を偏角分布として集計して、予め定めたC種の偏角区分の偏角の範囲を記録した偏角区分表を作成する偏角分布集計手段と、

前記部分ベクトル算定手段が作成したm個の部分ベクトル $v_1 \sim v_m$ のうち、部分空間bに対する部分ベクトル v_b ($b = 1 \sim m$)に対して、前記ノルム区分表を参照して前記各部分ベクトル v_b のノルムが所属するノルム区分の番号rを算定するノルム区分番号算定手段と、

前記部分ベクトル v_b と前記領域番号算定手段が算定した領域番号dの領域の中心方向を表す領域中心ベクトル p_d とのなす角の余弦である偏角($v_b \cdot p_d$) / ($|v_b| * |p_d|$)を算定して、前記偏角区分表を参照して、所属する偏角区分の番号cを算出する偏角区分番号算定手段と、

前記算定したノルム区分番号rに対応するノルム区分のノルムの最高値から、前記算定した部分ベクトル v_b の各成分 v_{bj} に対して、 v_{bj} が所属する予め定めた範囲の成分区分番号 w_j を算定する成分区分番号算定手段と、

前記部分空間番号bと、前記領域番号dと、前記偏角区分番号cと、前記ノルム区分番号rと、前記成分区分番号 w_j の列と、識別番号iとから、ベクトル索引に登録する索引登録データを算定する索引データ算定手段と、

前記ノルム区分表と、前記偏角区分表と、前記索引登録データとから、各部分ベクトルの識別番号と成分とを、部分空間番号b、領域番号d、偏角区分番号c、ノルム区分番号範囲 $[r_1, r_2]$ の組を鍵にして、検索可能で、かつ各ベクトルデータのベクトルの成分を、その識別番号で検索可能なベクトル索引として構

成する索引構成手段とを備えたことを特徴とするベクトル索引作成装置。

【請求項16】 前記部分ベクトル算定手段が、N次元ベクトルVの全成分を取り出すように、Vの先頭の成分から順に、N/m個、あるいは(N/m)+1個の成分を取り出して、部分ベクトルを作成する請求項14または15記載のベクトル索引作成装置。

【請求項17】 前記ノルム分布集計手段が、ノルム区分表を作成する際に、ノルム分布の集計結果をもとに、各ノルム区分に相当するノルム範囲に所属する部分ベクトルの数が、できる限り均一になるようにノルム区分を定める請求項14記載のベクトル索引作成装置。

【請求項18】 前記偏角分布集計手段が、偏角区分表を作成する際に、偏角分布の集計結果をもとに、各偏角区分に相当する偏角範囲に所属する部分ベクトルの数が、できる限り均一になるように偏角区分を定める請求項14記載のベクトル索引作成装置。

【請求項19】 前記領域番号算定手段が、部分ベクトルv_bの領域番号を、予め定めたD個の領域中心ベクトルp₁～p_Dのうち、p_dとv_bのなす角の余弦(v_b・p_d) / (|v_b| * |p_d|) がもっとも大きくなるような領域中心ベクトルp_dの番号dとして求める請求項14または15記載のベクトル索引作成装置。

【請求項20】 前記索引構成手段が、部分空間番号b、領域番号d、偏角区分番号c、ノルム区分番号rを結合した番号(b * Nd * Nc * Nr) + (d * Nc * Nr) + (c * Nr) + rを鍵とし、ベクトルの識別番号±と成分とを検索可能な探索木と、ベクトルデータの識別番号を添字とする、各部分ベクトルの前記探索木の鍵を記録した表とを作成し、ベクトル索引の一部とする請求項14から19のいずれかに記載のベクトル索引作成装置。

【請求項21】 前記領域番号算定手段が、成分が{-1, 0, +1}のうちのいずれかである、0ベクトルでない全てのベクトル(0, …, 0, +1)～(-1, …, -1)を正規化したベクトルを領域中心ベクトルとして用いる請求項14から20のいずれかに記載のベクトル索引作成装置。

【請求項22】 検索条件としてN次元実ベクトルの質問ベクトルQと、内積

下限値 α 、および最大取得ベクトル数Lとを指定して、少なくともN次元実ベクトルとそのID番号との組を有限個登録したベクトルデータから作成したベクトル索引を検索して、前記ベクトルデータベースのベクトルデータ(i, V)で、前記質問ベクトルQとの内積の値 $V \cdot Q$ が前記内積下限値 α より大きいものについて、その識別番号iと、QとVの内積の組(i, V · Q)を最大L個求める類似ベクトルの検索装置であって、

前記質問ベクトルQに対して、QのN個の成分を、前記ベクトル索引の作成時に用いた方法と同一の、予め定めた方法でm組に分けて、m個の部分質問ベクトル $q_1 \sim q_m$ を作成するとともに、各部分質問ベクトル q_b と対応する部分ベクトルとの内積（この内積をこれ以後「部分内積」と呼ぶ）の下限値である部分内積下限値 f_b を、指定された内積下限値 α から算定する部分質問条件算定手段と、

前記部分質問ベクトル q_b (b=1~m)と各領域bに対して、領域中心ベクトル p_d と前記部分質問ベクトル q_b との内積 $p_d \cdot q_b$ の値と、前記部分内積下限値 f_b と、前記ベクトル索引中のノルム区分表および偏角区分表とから、部分空間番号b、領域番号dにおいて検索すべき偏角区分番号c、ノルム区分の範囲 $[r_1, r_2]$ の組(c, [r₁, r₂])を算定する検索対象範囲生成手段と、

前記検索対象範囲生成手段が算定した(c, [r₁, r₂])を元に、(b, d, c, [r₁, r₂])を検索条件として前記ベクトル索引を範囲検索し、索引検索結果として、条件を満たす部分ベクトル v_b の識別番号iと成分とを求める索引検索手段と、

前記 v_b と q_b との部分内積 $v_b \cdot q_b$ と前記部分内積下限値 f_b との差である、部分内積差分 $(v_b \cdot q_b) - f_b$ を算出して、内積差分表の識別番号iの内積差分上限値S[i]として累算（加算）する内積差分上限算定手段と、

前記内積差分表S[i]中で値の大きいものから順に、前記ベクトル索引を識別番号iで検索してベクトルデータの成分Vを求め、Vと前記質問ベクトルQとの内積 $V \cdot Q$ から α を差し引いた内積差分値 $t = V \cdot Q - \alpha$ を算出していき、前記内積差分表中で、内積差分値を算出していない要素の最大値より大きい内積差分値を持つベクトルデータがL個以上になった時点、あるいは内積差分上限値が正である全てのベクトルデータの内積を算出した時点で、内積差分値の大きな最大

L個のベクトルデータに対し、少なくとも識別番号 i と内積 $t + \alpha$ との組を、検索結果として出力する類似検索結果決定手段とを備えたことを特徴とする類似ベクトル検索装置。

【請求項 23】 検索条件として N 次元実ベクトルの質問ベクトル Q と、距離上限値 α 、および最大取得ベクトル数 L を指定して、少なくとも N 次元実ベクトルとその識別番号との組を有限個登録したベクトルデータから作成したベクトル索引を検索して、前記質問ベクトル Q との内積の値が前記距離上限値 α 以下であるような、前記ベクトルデータ中の N 次元実ベクトル V の識別番号 i と、Q と V の距離 p の組 (i, p) を最大 L 個求める類似ベクトルの検索装置であって、

前記質問ベクトル Q に対して、Q の N 個の成分を、前記ベクトル索引の作成時に用いた方法と同一の、予め定めた方法で m 組に分けて、m 個の部分質問ベクトル $q_1 \sim q_m$ を作成するとともに、各部分質問ベクトル q_b と対応する部分ベクトル v_b との二乗距離 $|v_b - q_b|^2$ (すなわちユークリッド距離の二乗、この値のことをこれ以後「部分二乗距離」と呼ぶ) の上限値である部分二乗距離上限値 f_b を、指定された距離上限値 α から算定する部分質問条件算定手段と、

前記部分質問ベクトル q_b ($b = 1 \sim m$) に対して、前記部分質問ベクトル q_b と、前記部分二乗距離上限値 f_b と、前記ベクトル索引中のノルム区分表および偏角区分表とから、検索すべき部分空間番号 b、領域番号 d、偏角区分番号 c、ノルム区分の範囲 $[r_1, r_2]$ の組 (b, d, c, $[r_1, r_2]$) を系統的に生成する検索対象範囲生成手段と、

前記検索対象範囲生成手段が生成した (b, d, c, $[r_1, r_2]$) を検索条件として、前記ベクトル索引を範囲検索して、索引検索結果として、条件を満たす部分ベクトル v_b の識別番号 i と成分とを求める索引検索手段と、

前記部分二乗距離上限値 f_b と、 v_b と q_b との部分二乗距離 $|v_b - q_b|^2$ との差である、部分二乗距離差分 $f_b - |v_b - q_b|^2$ を算出して、二乗距離差分表の識別番号 i の二乗距離差分上限値 S[i] として累算 (加算) する二乗距離差分上限算定手段と、

前記二乗距離差分表 S[i] 中で値の大きいものから順に、前記ベクトル索引を識別番号 i で検索してベクトルデータの成分 V を求め、距離上限値の二乗 α^2 か

らVと前記質問ベクトルQとの二乗距離 $|V-Q|^2$ を差し引いた二乗距離差分値 $\alpha^2 - |V-Q|^2$ を算出していき、前記二乗距離差分表中で、二乗距離差分値を算出していない要素の最大値より大きい二乗距離差分値を持つベクトルデータがL個以上になった時点、あるいは二乗距離差分上限値が正である全てのベクトルデータの二乗距離差分値を算出した時点で、二乗距離差分値tの大きな最大L個のベクトルデータに対し、少なくとも識別番号iと距離 $(\alpha^2 - t)^{1/2}$ との組を、検索結果として出力する類似検索結果決定手段とを備えたことを特徴とする類似ベクトル検索装置。

【請求項24】 前記部分質問条件算定手段が、N次元ベクトルVの全成分を取り出すように、Vの先頭の成分から順に、N/m個、あるいは(N/m)+1個の成分を取り出して、部分質問ベクトルを作成する請求項22または23記載の類似ベクトル検索装置。

【請求項25】 前記部分質問ベクトル q_b と対応する部分ベクトル v_b との内積の下限値である部分内積下限値 f_b を、指定された内積下限値 α から、 $f_b = \alpha - |q_b|^2 / \sum (|q_b|^2)$ によって算定する請求項22記載の類似ベクトル検索装置。

【請求項26】 前記部分質問ベクトル q_b と対応する部分ベクトル v_b との二乗距離の上限値である部分二乗距離上限値 f_b を、指定された距離下上限値 α から、 $f_b = \alpha^2 - |q_b|^2 / \sum (|q_b|^2)$ によって算定する請求項23記載の類似ベクトル検索装置。

【請求項27】 請求項1から13のいずれかに記載の方法または請求項14から26のいずれかに記載の装置をソフトウェアにより実現したプログラムを記録した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、テキスト情報、画像情報、音声情報、アンケート結果、売り上げ(P.O.S)データなど、電子化されて蓄積された各種のデータベース(データの集合)中の各データの特徴抽出を行って作成したベクトルデータ(当該データの特徴

を表す実数をN個並べた、通常「特徴ベクトル」と呼ばれるN次元実ベクトルと、その識別番号などとの組)の集合であるベクトルデータベースに対して、計算機を利用してベクトルデータの検索、分類、傾向分析などを行うための索引の作成方法およびその装置と、指定したベクトルに類似のベクトルを、上記の方法および装置により作成された索引を用いて効率よく検索する、類似ベクトル検索方法およびその装置に関するものである。

【0002】

【従来の技術】

近年、テキスト、画像、音声などのマルチメディア情報のデータベース化や、POSシステムなどの普及に伴い、マルチメディアデータベースシステムやデータマイニングシステムなどのコンピュータシステムにおいて、数十～数百次元のベクトルデータを数十万個～数百万個集めたベクトルデータベースの検索、分類、傾向分析などを効率よく実行するための技術の研究開発が盛んになっている。

例えば新聞記事データベースの場合には、多数の新聞記事データを蓄積したデータベースに対し、W語の単語を集めた辞書を使って各新聞記事から辞書中の各単語kの出現回数 f_k を抽出し、各新聞記事を、識別番号iと、W次元の実ベクトル(f_1, f_2, \dots, f_W)の組で表現する。主成分分析手法などによってこのベクトルを変換し、各ベクトルの主要なN個($N \ll W$)の成分を求めて、これをベクトルデータとする。指定した新聞記事に対応するベクトルデータと、データベース中の他の新聞記事に対応するベクトルとの内積を計算し、内積の最も大きいベクトルを持つ新聞記事を求ることで、高精度な類似記事検索が可能になる。米国特許第4839853号には、このような、ベクトルデータを用いた文書検索方法が開示されている。

【0003】

また、写真データベースの場合には、多数の写真画像のデータを蓄積したデータベースに対し、各写真データに対して2次元のフーリエ変換を施し、主要なN個のフーリエ成分を f_k を抽出し、各写真データを、写真番号iと、N次元の実ベクトル(f_1, f_2, \dots, f_N)の組で表現し、これをベクトルデータとする。指定した写真に対応するベクトルデータと、データベース中の他の写真データ

タに対応するベクトルとの距離（2つのベクトルの差の大きさ）を計算し、距離の最も小さいベクトルを持つ写真データを求めてことで、高精度な類似写真検索が可能になる。さらに、例えば「ポートレート」、「風景写真」、「花の接写」といった、異なるカテゴリーに属する典型的な写真データをそれぞれ数枚ずつ、分類条件として提示し、各カテゴリーの平均的な特徴ベクトルを算出、各写真データのベクトルに対して最も距離の小さい特徴ベクトルのカテゴリーを割り当てることで、残りの写真データを上記の3種類のカテゴリーに自動的に分類することが可能になる。

【0004】

このような用途には、数十～数百次元という、極めて高次元のベクトルの効率のよい類似検索手法が必要となるため、各種の方法が研究されている。例えば、Norio Katayama and shinichi Satoh, "The SR-tree: An Index Structure for High-Dimensional Nearest Neighbor Queries", Proceedings of the SIGMOD '97, ACM (1997) には、SR木を用いた高次元ベクトルの索引作成方法および類似検索方法が開示されている。また、Sergey Brin, "Near Neighbor Search in Large Metric Spaces", Proceedings of the VLDB '95, Morgan-Kaufman Publishers (1995) には、ボロノイ分割に基いた高次元ベクトルの索引作成方法および類似検索方法が開示されている。さらに、Stefan Berchtold, Christian Bohm and Hans Kriegel, "The Pyramid-Technique: Towards Breaking the Curse of Dimensionality", Proceedings of the SIGMOD '98, ACM (1998) には、「ピラミッド技法」というデータ分割手法に基いた高次元ベクトルの索引作成方法および類似検索方法が開示されている。

【0005】

【発明が解決しようとする課題】

しかしながら、これらの従来のベクトル索引作成方法および類似ベクトル検索方法は、以下の4種の条件のいずれかが満たされず、広範囲の用途に広く適用できない、という課題があった。

- 1) ベクトルの次元が数百次元の場合でも、高速な検索ができる。
- 2) 類似検索の際に、ベクトル間の距離と、ベクトルの内積の2種類の類似度の

いずれかを選択できる。

3) 「もっとも類似しているL個のベクトルを求める」という型の類似検索ができる。さらに、Lが比較的大きい(数十~数百)場合でも、検索処理が極端に遅くならない。

4) 「内積が0.6以上」といった、類似検索範囲の指定ができる。

5) 索引作成に要する計算量が実用的範囲である(即ち、ベクトルデータ量nに比例した時間、あるいは $n * \log(n)$ 時間で索引が作成できる)。

具体的には、SR木を用いた方法では上記1) 2) が満たされず、ボロノイ分割に基いた方法では2), 5) が満たされず、ピラミッド技法を用いた方法では、2) 3) が満たされない。

【0006】

本発明のベクトル索引作成方法および類似ベクトル検索方法およびその装置は、これらの従来の技術の持つ課題を解決するもので、高次元のベクトルを、複数個の部分ベクトルに分解し、各部分ベクトルの方向と大きさを、中心ベクトルで規定した所属領域番号と、中心ベクトルとのなす角(偏角)と、ノルムを表すノルム区分との組で表現して記録することで、どのような質問ベクトルに対しても、ベクトル索引の検索対象範囲を精度よく限定することができ、部分内積の下限値(部分二乗距離の上限値)と、実際の部分内積(部分二乗距離)との差を累算することで、分岐限定技法による効率的な検索結果の確定ができるため、上記の1) ~4) をすべて満足した、広範囲の用途に適用可能なベクトル索引作成方法および類似ベクトル検索方法を提供するものである。

【0007】

【課題を解決するための手段】

上記課題を解決するために、本発明の第1の様態によるベクトル索引作成方法および装置は、部分ベクトルを算定する手段と、ノルム分布を集計してノルム区分表を作成する手段と、領域番号を算定する手段と、偏角分布を集計して偏角区分表を作成する手段と、ノルム区分番号を算定する手段と、偏角区分番号を算定する手段と、索引データを算定する手段と、索引を構成する手段とを有することにより、方向、ノルムの分布が不明なベクトルデータベースに対して、ベクトル

の次元が数百次元の場合でも、高速な検索が可能で、類似検索の際に、ベクトル間の距離と、ベクトルの内積の2種類の類似度のいずれかを選択でき、「もっとも類似しているL個のベクトルを求める」という型の類似検索ができ、さらに、Lが比較的大きい（数十～数百）場合でも、検索処理が極端に遅くならず、「内積が0.6以上」といった、類似検索範囲の指定が可能で、索引作成に要する計算量が実用的範囲であるような、ベクトル索引の作成を可能にする、という効果を奏するものである。

【0008】

また、本発明の第2の様態による、ベクトル索引作成方法および装置は、第1の様態に加えて、成分区分番号を算定する手段をさらに有することにより、第1の様態による効果に加えて、成分の量子化による計算誤差を最小限に抑えながら、作成すべきベクトル索引の容量を大幅に低減できる、という効果を奏するものである。

【0009】

また、本発明の第3の様態による、類似ベクトル検索方法および装置は、部分質問条件を算定する手段と、検索対象範囲を生成する手段と、索引を検索する手段と、内積差分上限を算定する手段と、類似検索結果を決定する手段とを有し、部分内積差分の累算値を算定して類似検索の手がかりとして用いることにより、ベクトルデータベースに対して、ベクトルの次元が数百次元の場合でも、高速な検索が可能で、「もっとも類似しているL個のベクトルを求める」という型の類似検索ができ、さらに、Lが比較的大きい（数十～数百）場合でも、検索処理が極端に遅くならず、「内積が0.6以上」といった、類似検索範囲の指定が可能で、内積を類似尺度として用いる類似ベクトル検索を可能にする、という効果を奏するものである。

【0010】

また、本発明の第4の様態による、類似ベクトル検索方法および装置は、部分質問条件を算定する手段と、検索対象範囲を生成する手段と、索引を検索する手段と、二乗距離差分上限を算定する手段と、類似検索結果を決定する手段とを有し、部分二乗距離差分の累算値を算定して類似検索の手がかりとして用いること

により、ベクトルデータベースに対して、ベクトルの次元が数百次元の場合でも、高速な検索が可能で、「もっとも類似しているL個のベクトルを求める」という型の類似検索ができ、さらに、Lが比較的大きい（数十～数百）場合でも、検索処理が極端に遅くならず、「距離が0.8以下」といった、類似検索範囲の指定が可能で、距離を類似尺度として用いる類似ベクトル検索を可能にする、という効果を奏するものである。

【0011】

【発明の実施の形態】

(実施の形態1)

以下、本発明の第1の実施の形態について、図面を参照しながら説明する。

(ベクトル索引作成装置の構成)

図1は本発明の請求項1、3～8、14、16～21にかかる、ベクトル索引作成装置の第1の実施の形態の全体構成を表わすブロック図である。図1において、ベクトルデータベース101は、20万件の新聞記事を集めた新聞記事全文データベースから作成した、各新聞記事の特徴を表す296次元の単位実ベクトルと、1から200000の範囲の識別番号の2項目からなるベクトルデータを20万個格納し、図12のような内容を持つ。

【0012】

部分ベクトル算定手段102は、ベクトルデータベース101中の各ベクトルデータの296次元のベクトルVに対し、37種の8次元部分ベクトル $v_0 \sim v_3$ ₆と、0～3-6の部分空間番号bとを算定する。

【0013】

ノルム分布集計手段103は、部分ベクトル算定手段102が20万個の各ベクトルデータについて算定した37個の各部分ベクトルのユークリッドノルムを算出して、その分布を集計し、256個の連続する実数範囲、

ノルム区分0 = [0, r1),

ノルム区分1 = [r1, r2),

...

ノルム区分255 = [r255, r256)

としてノルム区分を定める。

【0014】

ノルム区分表104は、ノルム分布集計手段103が算出したノルム区分を格納する。

【0015】

領域番号算定手段105は、部分ベクトル算定手段102が算定した8次元の各部分ベクトルvに対して、成分が{0, 1, -1}のいずれかであるような、0ベクトルでない8次元ベクトルをノルムが1になるよう正規化した、

領域中心ベクトル0 = (0, 0, 0, 0, 0, 0, 0, 1),

領域中心ベクトル1 = (0, 0, 0, 0, 0, 0, 0, -1),

領域中心ベクトル2 = (0, 0, 0, 0, 0, 0, 1, 0),

領域中心ベクトル3 = sqrt(1/2) * (0, 0, 0, 0, 0, 0, 0, 1),

領域中心ベクトル4 = sqrt(1/2) * (0, 0, 0, 0, 0, 0, 0, -1),

領域中心ベクトル5 = (0, 0, 0, 0, 0, 0, -1, 0),

...

領域中心ベクトル6554 = sqrt(1/7) * (-1, -1, -1, -1, -1, -1, 1, 0),

領域中心ベクトル6555 = sqrt(1/8) * (-1, -1, -1, -1, -1, -1, 1, 1),

領域中心ベクトル6556 = sqrt(1/8) * (-1, -1, -1, -1, -1, -1, 1, -1),

領域中心ベクトル6557 = sqrt(1/7) * (-1, -1, -1, -1, -1, -1, -1, 0),

領域中心ベクトル6558 = sqrt(1/8) * (-1, -1, -1, -1, -1, -1, -1, 1),

領域中心ベクトル6559 = sqrt(1/8) * (-1, -1, -1, -1, -1, -1, -1, -1),

という（ただし「sqrt(x)はxの平方根を表す」）、6560個のベクトルを領域中心ベクトルとし、部分ベクトルvとの内積が最も大きい領域中心ベクトル p_d を求めて、その番号dをvの所属領域の領域番号とし、 p_j とvとのなす角の余弦を、偏角cとして求める。

【0016】

偏角分布集計手段106は、領域番号算定手段105が20万個の各ベクトルデータの37個の部分ベクトルについて算定した偏角の値cの分布を集計し、4個の連続する実数範囲、

偏角区分0 = [c0, c1),

偏角区分1 = [c1, c2),

偏角区分2 = [c2, c3),

偏角区分3 = [c3, c4),

として偏角区分を定める。

【0017】

偏角区分表107は、偏角分布集計手段106算出した偏角区分を格納する。

【0018】

ノルム区分番号算定手段108は、部分ベクトル算定手段102が算出した各部分ベクトルのノルムの属するノルム区分番号rを、ノルム区分表104を検索して定める。

【0019】

偏角区分番号算定手段109は、部分ベクトル算定手段102が算出した各部分ベクトルvと、領域番号算定手段105がvに対して算定した領域中心ベクトルpから、vとpの偏角の属する偏角区分番号cを、偏角区分表107を検索して定める。

【0020】

索引データ算定手段110は、部分ベクトル算定手段102の算定した部分ベクトル v_b および部分空間番号bと、領域番号算定手段105の算定した領域番号dと、偏角区分番号算定手段109の算定した偏角区分番号cと、ノルム区分番号算定手段108の算定したノルム区分番号rとから検索用の鍵、

$$K = ((b * 6560 + d) * 4 + c) * 256 + r$$

を作成し、この鍵Kと部分ベクトルの識別番号iと成分 v_b の組(K, i, v_b)を索引データとして算定する。

【0021】

索引構成手段111は、索引データ算定手段110の算定した索引データ(K, i, v_b)から、Kを鍵とし、(i, v_b)を検索するような探索木と、各識別番号iと各部分空間番号bの組に対して、領域番号dと偏角区分番号cと、ノルム区分番号rとから第2の鍵、

$$L = (d * 4 + c) * 256 + r$$

を格納した逆探索表と、ノルム区分表104と、偏角区分表107とを格納した索引を構成する。

【0022】

ベクトル索引112は、索引構成手段111が作成する探索木と逆探索表とノルム区分表104と偏角区分表107とを格納する。

【0023】

(ベクトル索引作成装置の動作)

以上の構成のベクトル索引作成装置において、その動作を図面をもとに説明する。図5はベクトル索引作成の第1ステップにおけるノルム区分表Rおよび偏角区分表Cの作成処理の手順を記述した流れ図であり、図6はベクトル索引作成の第2、第3ステップにおける索引登録データの算定およびベクトル索引の作成処理の手順を記述した流れ図である。図中の「sqrt(x)」はxの平方根を、「int(x)」はxの整数部分を、「abs(x)」はxの絶対値を、それぞれ表すものとする。また「sign(x)」はxが非負なら1を、負なら2をとる関数であるとする。

【0024】

(ベクトル索引作成の第1ステップ)

ベクトル索引作成の第1ステップでは、まず、部分ベクトル算定手段102が、ベクトルデータベース101からベクトルデータを順に読み込んで部分ベクトルを算定し、ノルム分布集計手段103と偏角分布集計手段106が、それぞれ

部分ベクトルのノルム分布、偏角分布を算定する。全てのベクトルデータを処理した時点で、ノルム区分表および偏角区分表が作成される。ベクトルデータベース中のベクトルのノルムの上限値は既知であり、上限値は r_{sup} であるとする。本実施の形態の例では、各ベクトルデータのベクトルが単位ベクトルなので、明らかに $r_{sup} = 1$ となる。ベクトルデータベース中のベクトルのノルムの上限値が未知である場合には、予め調査して r_{sup} を求めておけばよい。

【0025】

初めに、ステップ1001で、集計用の表 H_r および H_c を0に初期化し、総部分ベクトル数 n も0にする。次に、ステップ1002で、未処理のベクトルデータ (i, v) をベクトルデータベースから1つ読み込む。部分空間の番号 b を0に初期化する。ステップ1003で、 b の値に従って、8次元の部分ベクトル u を、読み込んだ296次元のベクトル v の先頭から、連続する8成分ずつに区分して37種作成する。例えば図12の1番目のベクトルデータの場合には、 $b = 0$ の部分ベクトルは、

$$(+0.029259 -0.016005 -0.021118 +0.024992 -0.006860 -0.009032 -0.007255
-0.007715)$$

となる。 $b = 1$ の部分ベクトルは、

$$(-0.025648 +0.016061 -0.060584 -0.013593 -0.020985 -0.112403 -0.012045
+0.044741)$$

となり、 $b = 36$ の部分ベクトルは、

$$(+0.069379 +0.020206 +0.032996 +0.047815 +0.046106 +0.001794 +0.035342
-0.003895)$$

となる。そして、 u のノルム $|u|$ をノルムの最大値 r_{sup} で割って100倍した後、整数化し、ノルム分布集計表 H_r の該当する区分 j に累算し、ノルム分布を集計する。

【0026】

図13はこのようにして集計したノルム分布のグラフの例である。グラフの横軸はノルム分布集計表 H_r の区分番号、縦軸は各区分番号 j に対する $H_r[j]$ の値、即ち区分 j のノルム範囲のノルムを持つ部分ベクトルの数である。図12

の1番目のベクトルデータの $b = 0$ の部分ベクトルの場合には、

$|u| = \sqrt{0.029259*0.029259 + 0.016005*0.016005 + \dots + 0.007715*0.007715} = 0.049193$

であり、 $r_s u p = 1$ であるから、区分 j は、

$j = \text{int}(0.049193 / 1.0) * 10000 = 491$

となる。

【0027】

偏角区分は、ステップ1004～ステップ1009で集計する。まず、ステップ1004で、部分ベクトル u の8個の成分 $u[0] \sim u[7]$ に対して、絶対値の大きい順に、成分の番号を格納する。図12の1番目のベクトルデータの $b = 0$ の部分ベクトルの場合には、第0成分の絶対値が最も大きく、次に第3成分の絶対値が大きく、第4成分の絶対値が最も小さいので、

$s[0..7] = (0 3 2 1 5 7 6 4)$

となる。

【0028】

次にステップ1005～ステップ1008を、変数 m の値を0から7まで変えながら、8回（8=部分空間の次元）繰り返し、6560種の領域中心ベクトルのうちで最も部分ベクトル u との内積が大きいものの番号 d と、そのときの内積の値 x を求める。ステップ1005では、絶対値の大きいほうから $m+1$ 個の成分が、（部分ベクトルの成分の符号） * 1 で、残りの $7-m$ 個の成分が0であるような領域中心ベクトルの番号 j と内積の \sqrt{m} 倍の値 y を求めている。ステップ1006では1005で求めた値 y から内積を $y * \sqrt{1/m}$ で計算し、これまでの内積の最高値 x を比較し、 x より大きければステップ1007で内積の最高値 x と、領域中心ベクトルの番号 d を更新する。このように、成分が $\{+1, -1\}$ のいずれかであるような領域中心ベクトル群を用いることにより、非常に簡単な計算によって、部分ベクトルと最も内積の大きい領域中心ベクトルの番号と、その内積の値とを効率よく求めることができる。

【0029】

図12の1番目のベクトルデータの $b = 0$ の部分ベクトルの場合には、

$(|u[0]|) * \sqrt{1/1} = 0.029259$
 $(|u[0]| + |u[3]|) * \sqrt{1/2} = 0.038361$
 $(|u[0]| + |u[3]| + |u[2]|) * \sqrt{1/3} = 0.043514$
 $(|u[0]| + |u[3]| + |u[2]| + |u[1]|) * \sqrt{1/4} = 0.045687$
 $(|u[0]| + |u[3]| + |u[2]| + |u[1]| + |u[5]|) * \sqrt{1/5} = 0.044903$
 $(|u[0]| + |u[3]| + |u[2]| + |u[1]| + |u[5]| + |u[7]|) * \sqrt{1/6} = 0.044140$
 $(|u[0]| + |u[3]| + |u[2]| + |u[1]| + |u[5]| + |u[7]| + |u[6]|) * \sqrt{1/7} =$
 0.043608
 $(|u[0]| + |u[3]| + |u[2]| + |u[1]| + |u[5]| + |u[7]| + |u[6]| + |u[4]|) * \sqrt{1/8} = 0.043217$

となって、内積の最高値である $x = 0.045687$ 、および領域中心ベクトル
 $(+1/2, -1/2, -1/2, +1/2, 0, 0, 0, 0)$

の番号である $d = (3^7) + 2 * (3^6) + 2 * (3^5) + (3^4) = 4212$ が求まる。

【0030】

続いてステップ1009で内積 x を部分ベクトル u のノルムで割って、部分ベクトルと領域中心ベクトルとのなす角の余弦を求め、10000倍した後、整数化し、偏角分布集計表 Hc の該当する区分 j に累算し、偏角分布を集計する。図14このようにして集計した偏角分布のグラフの例である。グラフの横軸は偏角分布集計表 Hc の区分番号、縦軸は各区分番号 j に対する $Hc[j]$ の値、即ち区分 j の偏角範囲の偏角を持つ部分ベクトルの数である。ただし図14において、8274より小さな区分の Hc の集計値は全て0なので、8000～10000区分の部分のみ図示している。図12の1番目のベクトルデータの $b=0$ の部分ベクトルの場合には、

```
j = int(10000 * 0.045687 / 0.049193)
= int(10000 * 0.928730) = 9287
```

となる。

【0031】

部分ベクトル選択用変数 b と、総部分ベクトル数集計用変数 n を増やした後、

ステップ1010で着目ベクトルデータの全ての部分ベクトルを処理したか否かを判定し、まだ未処理の部分ベクトルが残っている場合にはステップ1003に戻って、次の部分ベクトルに対して処理を行う。全ての部分ベクトルを処理した場合には、ステップ1011でベクトルデータベース101中の全てのベクトルデータを処理したか否かを判定し、まだ未処理のベクトルデータが残っている場合にはステップ1002に戻って次のベクトルデータを処理し、全てのベクトルデータを読み込んで処理した場合には、ステップ1012～ステップ1018のノルム区分表、偏角区分表の作成処理に進む。

【0032】

ステップ1012では作業変数を初期化し、ステップ1013～ステップ1018でノルム区分表および偏角区分表の区分データの作成処理を行う。ステップ1013では、ノルム集計結果のうち、ノルムが0から $r_{sup} * j / 1000$ までの部分ベクトルの個数の合計値xと、偏角集計結果のうち、偏角が0から $j / 10000$ までの部分ベクトルの個数の合計値yとを求める。

【0033】

ステップ1014では、ノルムが0から $r_{sup} * j / 10000$ までの部分ベクトルの個数の全部分ベクトル数に対する割合 x / n が、ノルム区分表の256個の区分のうち、第k区分までの区分数の割合 $k / 256$ より大きいか否かを判定する。大きい場合には、ステップ1015に進んでノルム区分表の第k区分の境界値である $R[k]$ を $r_{sup} * j / 10000$ と定める。図15は、以上のようにして図13のノルム分布のノルム分布集計表H_rから作成したノルム区分表の例である。分布が集中している0.1～0.2の区間を細かく区分していることがわかる。

【0034】

ステップ1016からステップ1017で、偏角区分についても、同様にして偏角区分表の第m区分の境界値を定める。ステップ1018でノルム集計結果、偏角集計結果をすべて処理したか否かを判定し、未処理の集計結果が残っている場合にはステップ1013に戻って処理を続行し、全ての集計結果を処理し終えた場合にはステップ1019に進んで、R[0..256], C[0..4]をそれぞれノル

ム区分表、偏角区分表としてベクトル索引作成の第1ステップを終了する。図16は、以上のようにして図14の偏角分布の偏角分布集計表Hcから作成した偏角区分表の例である。分布が集中している0.95付近を細かく区分していることがわかる。

【0035】

(ベクトル索引作成の第2ステップ)

ベクトル索引作成の第2ステップでは、図6のステップ1101～ステップ1109に記載した処理を行って、個々の部分ベクトルから索引登録データを作成する。まず、1101で探索木Tを初期化し、Tの登録データ数を0とする。探索木としては、

1) 整数値を鍵として、ベクトルデータ(i, u)即ち整数と8個の浮動小数点数の組が登録できる。

2) 登録時の整数値の範囲を鍵にして、登録したデータを検索できる。

という、2つの条件を満たすものであれば、R. セジウィック著、野下浩平他訳、"アルゴリズム第2巻 探索・文字列・計算幾何", 近代科学社(1992)や、G. H. Gonnet著 玄光男他訳 "アルゴリズムとデータ構造ハンドブック", 啓学出版(1987)などの教科書に記載されている、B木や2分探索木などの(平衡)探索木を用いることができる。

【0036】

ステップ1102では、ベクトルデータベース101から未処理のベクトルデータを1つ読み込み、部分空間番号bを0から順に増やしながら各部分空間の部分ベクトルに対する処理を行う。ステップ1103では、部分ベクトルuを作成し、作成済みのノルム区分表104を検索して、ノルム|u|に対応するノルム区分の番号rを求める。ステップ1104～ステップ1108までは、図5のステップ1004～ステップ1008と同一の処理を行って、6560種の領域重心ベクトルのうちで最も部分ベクトルuとの内積が大きいものの番号dと、そのときの内積の値xとを求める。

【0037】

ステップ1109では、作成済みの偏角区分表107を検索して、偏角(即ち

部分ベクトルと所属領域の領域中心ベクトルとのなす角の余弦) $x / |u|$ に
対応する偏角区分の番号 c を求める。ステップ1110では、索引データ算定手
段110が、こうして求めたノルム区分番号 d 、偏角区分番号 c から、

$$\begin{aligned} k &= b * N_d * N_c * N_r + d * N_c * N_r + c * N_r + r \\ &= b * 7617440 + d * 1024 + c * 256 + r \end{aligned}$$

によって、部分空間番号 b 、領域番号 d 、偏角区分番号 c 、ノルム区分番号 r の
4種の整数値を1つの整数値に変換し、探索木に登録する際の鍵 k を算定、ステ
ップ1111でこの鍵 k と、部分ベクトルデータ (i, u) から、索引登録データ
(k, i, u) を算定する。ただし、 N_d は領域の総数で6560、 N_c は偏角
区分数で4、 N_r はノルム区分数で256である。このようにして、ベクトル索
引作成の第2ステップで、各ベクトルデータの各部分ベクトルに対する索引登録
データ (k, i, u) を効率よく（ベクトルデータ数に比例した時間で）作成で
きる。

【0038】

（ベクトル索引作成の第3ステップ）

ベクトル索引作成の第3ステップでは、図6のステップ1111～ステップ1
115に記載した処理を行って、索引登録データからベクトル索引を作成する。
まず、ステップ1111では、索引登録データ (k, i, u) 中の k を鍵として
、データ (i, u) を探索木に（追加）登録する。次にステップ1112で、逆
探索表Kの識別番号 i のベクトルデータの部分空間番号 b に対応する要素K [i, u]
に鍵 k を格納する。部分空間番号 b を1増やした後、ステップ1113で全て
の部分空間について処理を終了したか否かを判定する。まだ未処理の部分空間が
残っている場合にはステップ1103に戻って次の部分ベクトルに対する処理を行
い、全ての部分空間について処理を終了した場合には、ステップ1114に進
む。ステップ1114では、ベクトルデータベース101中の全てのベクトルデ
ータを処理したか否かを判定する。まだ未処理のベクトルデータが残っている場
合にはステップ1102に戻って次のベクトルデータに対する処理を行い、全
てのベクトルデータについて処理を終了した場合には、ステップ1105に進んで
探索木T、逆探索表K、ノルム区分表R、偏角区分表Cを格納したベクトル索引

を作成して、ベクトル索引の作成が完了する。

【0039】

以上説明したように、本発明の第1の実施の形態のベクトル索引作成方法および装置によれば、以下のような優れた効果を奏する。

1) 296次元のベクトルを、37種の8次元の部分ベクトルに分解し、各々の部分ベクトルに対して、その方向を、6560種の領域中の所属領域の領域番号と偏角区分番号の組で精密に量子化し、その大きさをノルム区分番号で量子化した上で、複数の鍵を1つの整数値にエンコードして探索木に登録することで、部分空間毎の高速で高精度な範囲検索が可能となる。

2) また、逆探索表を作成・装備することで、ベクトルデータの識別番号を指定してベクトルの成分を取得する機能が、成分データを2重持ちすることなく実現できるため、元のベクトルデータベース101が検索時には不要となり、検索装置の記憶容量を削減することが可能となる。

3) ノルム分布集計手段、偏角分布集計手段において、各区分に所属する部分ベクトルの個数ができる限り均一になるように区分境界を定めるため、分布に偏りのあるベクトルデータベースに対しても、常に最適な（検索速度の低下を最小限に抑えた）ベクトル索引を作成することができる。

4) 領域中心ベクトルとして、成分が(0, +1, -1)のいずれかである0ベクトルでない全てのベクトルを正規化したベクトルの集合を用いているため、各部分ベクトルの所属領域の算定が、領域の個数には依存せずに見え、部分ベクトルの成分の絶対値順序の算出と、成分の絶対値の加算などの、極めて少ない計算量で済むことから、数十万～数百万個のベクトルデータからなる大規模なベクトルデータベースの場合でも、実用的な処理時間でベクトル索引を作成することが可能となる。

【0040】

(実施の形態2)

次に、本発明の第2の形態について、図面を参照しながら説明する。

(ベクトル索引作成装置の構成)

図2は本発明の請求項2、3～8、15、16～21にかかる、ベクトル索引

作成装置の第2の実施の形態の全体構成を表わすブロック図である。図2において、ベクトルデータベース201は、20万件の新聞記事を集めた新聞記事全文データベースから作成した、各新聞記事の特徴を表す296次元の単位実ベクトルと、1から200000の範囲の識別番号と、記事の見出しの3項目からなるベクトルデータを20万個格納し、図12のような内容を持つ。

【0041】

部分ベクトル算定手段202は、ベクトルデータベース201中の各ベクトルデータの296次元のベクトルVに対し、37種の8次元部分ベクトル $v_0 \sim v_3$ ₆と、0～36の部分空間番号bとを算定する。

【0042】

ノルム分布集計手段203は、部分ベクトル算定手段202が20万個の各ベクトルデータについて算定した37個の各部分ベクトルのユークリッドノルムを算出して、その分布を集計し、256個の連続する実数範囲、

ノルム区分0 = [0, r1],

ノルム区分1 = [r1, r2],

...

ノルム区分255 = [r255, r256],

としてノルム区分を定める。

【0043】

ノルム区分表204は、ノルム分布集計手段203が算出したノルム区分を格納する。

【0044】

領域番号算定手段205は、部分ベクトル算定手段202が算定した8次元の各部分ベクトルvに対して、成分が{0, 1, -1}のいずれかであるような、0ベクトルでない8次元ベクトルをノルムが1になるよう正規化した、

領域中心ベクトル0 = (0, 0, 0, 0, 0, 0, 0, 1),

領域中心ベクトル1 = (0, 0, 0, 0, 0, 0, 0, -1),

領域中心ベクトル2 = (0, 0, 0, 0, 0, 0, 1, 0),

領域中心ベクトル3 = sqrt(1/2) * (0, 0, 0, 0, 0, 0, 1,

1) ,

領域中心ベクトル4 = sqrt(1/2) * (0, 0, 0, 0, 0, 0, 1,
-1) ,

領域中心ベクトル5 = (0, 0, 0, 0, 0, 0, -1, 0) ,

...

領域中心ベクトル6554 = sqrt(1/7) * (-1, -1, -1, -1
, -1, -1, 1, 0) ,

領域中心ベクトル6555 = sqrt(1/8) * (-1, -1, -1, -1
, -1, -1, 1, 1) ,

領域中心ベクトル6556 = sqrt(1/8) * (-1, -1, -1, -1
, -1, -1, 1, -1) ,

領域中心ベクトル6557 = sqrt(1/7) * (-1, -1, -1, -1
, -1, -1, -1, 0) ,

領域中心ベクトル6558 = sqrt(1/8) * (-1, -1, -1, -1
, -1, -1, -1, 1) ,

領域中心ベクトル6559 = sqrt(1/8) * (-1, -1, -1, -1
, -1, -1, -1, -1) ,

という（ただし「sqrt(x) はxの平方根を表す」）、6560個のベクトルを領域中心ベクトルとし、部分ベクトルvとの内積が最も大きい領域中心ベクトル p_d を求めて、その番号dをvの所属領域の領域番号とし、 p_j とvとのなす角の余弦を、偏角cとして求める。

【0045】

偏角分布集計手段206は、領域番号算定手段205が20万個の各ベクトルデータの37個の部分ベクトルについて算定した偏角の値cの分布を集計し、4個の連続する実数範囲。

偏角区分0 = [c0, c1) ,

偏角区分1 = [c1, c2) ,

偏角区分2 = [c2, c3) ,

偏角区分3 = [c3, c4) ,

として偏角区分を定める。

【0046】

偏角区分表207は、偏角分布集計手段206算出した偏角区分を格納する。

【0047】

ノルム区分番号算定手段208は、部分ベクトル算定手段202が算出した各部分ベクトルのノルムの属するノルム区分番号rを、ノルム区分表204を検索して定める。

【0048】

偏角区分番号算定手段209は、部分ベクトル算定手段202が算出した各部分ベクトルvと、領域番号算定手段205がvに対して算定した領域中心ベクトルpから、vとpの偏角の属する偏角区分番号cを、偏角区分表207を検索して定める。

【0049】

索引データ算定手段210は、部分ベクトル算定手段202の算定した部分ベクトルvbおよび部分空間番号bと、領域番号算定手段205の算定した領域番号dと、偏角区分番号算定手段209の算定した偏角区分番号cと、ノルム区分番号算定手段208の算定したノルム区分番号rとから検索用の鍵、

$$K = ((b * 6560 + d) * 4 + c) * 256 + r$$

を作成し、この鍵Kと部分ベクトルの識別番号iと成分区分番号yjの組(K, i, y)を索引データとして算定する。

【0050】

索引構成手段211は、索引データ算定手段210の算定した索引データ(K, i, y)から、Kを鍵とし、(i, y)を検索するような探索木と、各識別番号iと各部分空間番号bの組に対して、領域番号dと偏角区分番号cと、ノルム区分番号rとから第2の鍵、

$$L = (d * 4 + c) * 256 + r$$

を格納した逆探索表と、ノルム区分表204と、偏角区分表207とを格納した索引を構成する。

【0051】

ベクトル索引212は、索引構成手段211が作成する探索木と逆探索表とノルム区分表204と偏角区分表207とを格納する。なお、構成要素201～212は、それぞれ図5の構成要素101～112と対応し、特に構成要素201～209は、図5の構成要素101～109と同一である。

【0052】

成分区分番号算定手段213は、部分ベクトル算定手段202が算定した部分ベクトル v_b と、ノルム区分番号算定手段208が算定したノルム区分番号とから、部分ベクトルの各成分の値から、0～255の範囲の成分区分番号 $y_0 \sim y_7$ を算定する。

【0053】

(ベクトル索引作成装置の動作)

(ベクトル索引作成の第1ステップ)

以上の構成のベクトル索引作成装置において、その動作を図面をもとに説明する。ベクトル索引作成の第1ステップにおけるノルム区分表Rおよび偏角区分表Cの作成処理の手順は、第1の実施の形態における手順と同一であり、同一のベクトルデータベースならば、作成されるノルム区分表Rおよび偏角区分表Cは、ともに第1の実施の形態におけるノルム区分表Rおよび偏角区分表Cの内容と同一となるので、説明を省略する。

【0054】

(ベクトル索引作成の第2、第3ステップ)

図7はベクトル索引作成の第2、第3ステップにおける索引登録データの算定およびベクトル索引の作成処理の手順を記述した流れ図である。図7のステップ1200～ステップ1216はそれぞれ図6のステップ1100～ステップ1116に対応し、特に、ステップ1211、ステップ1215、ステップ1217以外の各ステップは、図6の対応するステップと同一の処理であるため、説明を省略する。

【0055】

ステップ1217では、部分ベクトル $u[0..7]$ から、 u の各成分に対応する成分区分番号 $y[0..7]$ を算定する。どの $u[m]$ に対しても、 $\text{abs}(u[m]) \leq |u| < R$

$r+1]$ なので、

$$-1 < u[m] / R[r+1] < +1$$

が成り立ち、成分区分番号 $y[m]$ は 0 ~ 255 までの、8 ビットで表現できる整数値となる。ステップ 1211 では、 u の代わりに、 y を使って、(i, y) という、整数データを k を鍵にして探索木 T に登録する。各々の $y[m]$ は 8 ビットで表現できるので、 $u[m]$ を浮動小数点形式で登録する場合に比べ、探索木 T の容量が大幅に低減する。ステップ 1215 では、このようにして作成した探索木 T を含むベクトル索引を作成するので、結果的に作成したベクトル索引の容量が、 $u[m]$ を登録する場合に比べて小さくて済む。

【0056】

なお、本実施の様態 2においては、ステップ 1217 で 8 ビットの整数値 $y[m]$ で各成分 $u[m]$ を近似したが、8 ビットでは類似検索時の精度が不足する場合には、十分な精度が得られるよう、9 ビット～24 ビットで表現して登録すればよい。

【0057】

以上説明したように、本発明の第 2 の実施の形態のベクトル索引作成方法および装置によれば、以下のような優れた効果を奏する。

1) 296 次元のベクトルを、37 種の 8 次元の部分ベクトルに分解し、各々の部分ベクトルに対して、その方向を、6560 種の領域中の所属領域の領域番号と偏角区分番号の組で精密に量子化し、その大きさをノルム区分番号で量子化し、さらに加えて部分ベクトルの各成分を成分区分番号という、ノルム区分を基準にした量子化を行って、複数の鍵を 1 つの整数値にエンコードして、近似結果である部分ベクトルの成分区分番号とともに探索木に登録することで、部分空間毎の高速で高精度な範囲検索が可能となる。

2) 逆探索表を作成・装備することで、ベクトルデータの識別番号を指定してベクトルの成分を取得する機能が、成分データを 2 重持ちすることなく実現できるため、元のベクトルデータベース 101 が検索時には不要となり、検索装置の記憶容量を削減することが可能となる。

3) ノルム分布集計手段、偏角分布集計手段において、各区分に所属する部分

ベクトルの個数ができる限り均一になるように区分境界を定めるため、分布に偏りのあるベクトルデータベースに対しても、常に最適な（検索速度の低下を最小限に抑えた）ベクトル索引を作成することができる。

4) 領域中心ベクトルとして、成分が {0, +1, -1} のいずれかである0ベクトルでない全てのベクトルを正規化したベクトルの集合を用いているため、各部分ベクトルの所属領域の算定が、領域の個数には依存せずに行え、部分ベクトルの成分の絶対値順序の算出と、成分の絶対値の加算などの、極めて少ない計算量で済むことから、数十万～数百万個のベクトルデータからなる大規模なベクトルデータベースの場合でも、実用的な処理時間でベクトル索引を作成することが可能となる。

5) 作成するベクトル索引の容量を大幅に削減することができる。

【0058】

(実施の形態3)

次に、本発明の第3の形態について、図面を参照しながら説明する。

(類似ベクトル検索装置の構成)

図3は本発明の請求項9、11、12、22、24、25にかかる、類似ベクトル検索装置の全体構成を表わすブロック図である。図3において、ベクトル索引301は、上記した第1の実施の形態のベクトル索引作成装置によって作成したものであり、20万件の新聞記事を集めた新聞記事全文データベースから作成した、各新聞記事の特徴を表す296次元の実ベクトルと、各記事を一意に識別するための1から200000の範囲の識別番号の2項目からなるベクトルデータを20万個格納した、図12のような内容を持つベクトルデータベースから作成したベクトル索引である。

【0059】

検索条件入力手段3-0-2は、前記新聞記事全文データベースに対する類似検索を行うために、前記新聞記事全文データベース中のいずれかの記事の識別番号と、類似検索範囲を示す0～100までの類似度下限値と最大取得件数とを入力し、ベクトル索引301を識別番号で検索して入力された識別番号から対応する記事のベクトルを質問ベクトルQとして求め、類似度下限値から内積下限値 α を求め

る。

【0060】

部分質問条件算定手段303は、検索条件入力手段302が求めた質問ベクトルQに対し、0～36までの部分空間について、37種の8次元部分質問ベクトルqと、qに対応する部分ベクトルとの内積の下限値である部分内積下限値fを $f = \alpha |q|^2 / |Q|^2$ によって算定する。

【0061】

検索対象範囲生成手段304は、部分質問条件算定手段303が求めた部分空間bに対する部分質問ベクトルqと部分内積下限値fとベクトル索引301中のノルム区分表および偏角区分表とから、部分質問ベクトルqとの部分内積が部分内積下限値fより大きい可能性がある部分文書ベクトルを含んでいる領域を特定する領域番号d、偏角区分番号c、ノルム区分範囲 $[r_1, r_2]$ の組(d, c, $[r_1, r_2]$)を全て列挙する。

【0062】

索引検索手段305は、検索対象領域生成手段304が各部分空間bに対して生成した(d, c, $[r_1, r_2]$)から、ベクトル索引301に対する検索条件Kを、ベクトル索引作成時における鍵の算定と同様に、

$$K = [k_{\min}, k_{\max}]$$

$$k_{\min} = b * 7617440 + d * 1024 + c * 256 + r_1$$

$$k_{\max} = b * 7617440 + d * 1024 + c * 256 + r_2$$

と算定し、ベクトル索引301を検索条件Kで範囲検索して検索条件に合致する鍵を持つ部分ベクトルvと識別番号iとの組(i, v)をすべて求める。

【0063】

内積差分上限算定手段306は、索引検索手段305が求めた部分ベクトルvと識別番号iとの組(i, v)および部分質問条件算定手段303が求めた部分質問ベクトルqと部分内積下限値fから、 $t = (v \cdot q) - f$ によって部分内積差分値tを算定して、識別番号iを添字とする表の要素S[i]に、部分内積差分値tを累算(加算)することで、識別番号iのベクトルデータのベクトルVと質問ベクトルQとの内積Q・Vから内積下限値 α を減じた内積差分の上限値を算

定する。

【0064】

内積差分表307は、内積差分上限算定手段306が算出した内積差分の上限値を蓄積するものであり、識別番号*i*のベクトルベータの内積差分値S[i]を参照／格納する。

【0065】

類似結果決定手段308は、内積差分表307の要素S[i]中で、内積差分上限値S[i]が正の値で大きいものから順に、ベクトル索引301を識別番号*i*で検索して対応するベクトルVを求め、Vと検索条件入力手段302が算出した質問ベクトルQとの内積V・Qから、検索条件入力手段302が算出した内積下限値 α を差し引いた内積差分値V・Q - α を算出して、S[i]を内積差分値V・Q - α で置き換え、内積差分値を算出していない記事の部分内積差分累算値の最大値より大きい内積差分値を持つ、内積差分値を算出した記事がL個以上になった時点、あるいは部分内積差分累算値が正である全ての記事の内積差分値を算出した時点で、内積差分値が正である、内積差分値の大きな最大L個の結果候補(i, S[i])に対し、識別番号*i*と内積S[i] + α との組(i, S[i] + α)を、検索結果として、検索結果出力手段309に出力する。

【0066】

検索結果出力手段309は、類似ベクトル検索の結果である、最大L件の新聞記事の識別番号と0～100の範囲の類似度とを、類似結果決定手段308が求めた検索結果から算定して表示する。

【0067】

(類似ベクトル検索装置の動作)

以上の構成の類似ベクトル検索装置において、その動作を図面をもとに説明する。図8は類似ベクトル検索の第1ステップにおける検索処理の手順を記述した流れ図、図9は類似ベクトル検索の第2ステップにおける検索処理の手順を記述した流れ図である。類似ベクトル検索の第1ステップでは、検索条件入力手段302から入力された検索条件から部分質問ベクトルqと部分内積下限値fを作成し、ベクトル索引301を検索して、各ベクトルデータの内積差分上限値S[

i]、即ち質問ベクトルとの内積から内積下限値を減じた値がS[i]未満となるような値を内積差分表307に求める。次に、類似ベクトル検索の第2ステップでは、第1ステップで内積差分表307に求めた内積差分上限値を手がかりにして、類似検索結果決定手段308が、「質問ベクトルとの内積が α より大きい」という検索条件に合致し、かつ質問ベクトルとの内積が比較的大きいベクトルデータから順に、ベクトルの成分を検索して内積差分を求めていき、内積差分を求めていないベクトルデータのどれよりも内積差分値が大きいことが保証されるベクトルデータが指定件数（即ちL件）以上集まるか、あるいは検索条件に合致する全てのベクトルデータについて内積差分値を求めるまで処理を続け、得られた内積差分値から内積を算出して最終結果を出力する。

【0068】

（類似ベクトル検索の第1ステップ）

以下、検索条件として識別番号1、類似度下限値90、最大取得件数10が入力された場合を例にとって、図8、図9にそって類似ベクトル検索の内容を説明する。識別番号が1なので、296次元のベクトルの各成分は、図12の上半分の通りとなる。まず、ステップ1301で内積差分表Sの20万個の要素S[0]～S[200000]を0に初期設定する。次に、上記の検索条件を検索条件入力手段302から読み込み、それぞれi, Z, Lに格納する。

【0069】

ステップ1302で部分空間番号bを0に初期化した後、類似度下限値Zから、内積下限値 α を算出する。この検索条件の場合には、 $\alpha \leftarrow (90 - 50) / 5 = 0.8$ となる。ステップ1304、ステップ1305で、各部分空間について、ベクトル索引301の逆変換表Kを引いて鍵を求め、探索表を検索してベクトルデータを求め、識別番号が1であるもののベクトル部分をQに格納することで、質問ベクトルをQ[0..295]に求める。ステップ1306で部分空間番号を初期化した後、ステップ1307～ステップ1317で各部分空間についてベクトル索引を検索して内積差分表307に各ベクトルデータの内積差分上限値を求める。

【0070】

ステップ1307では、部分空間番号bの部分質問ベクトルq [0..7]と部分内積下限値f、即ちこの部分空間の部分ベクトルデータとqとの内積の下限値を求める。b = 0の場合には、 $|q|^2 = 0.221795$, $|Q|^2 = 1$ であるから、

$$f = 0.8 * 0.221795 / 1.0 = 0.177436$$

となる。領域番号dを0に初期化したあと、検索対象範囲を決定するために用いる表Wを作成する。表Wは偏角区分番号c、ノルム区分番号rで引いて、着目している領域番号dの領域の中心ベクトルpと部分質問ベクトルqとの内積 $p \cdot q$ がW [c, r]未満ならば、区分(d, c, 0)～区分(d, c, r)の部分ベクトルvと部分質問ベクトルqとの内積がf以下となるように作成する。この場合には、区分(d, c, 0)～区分(d, c, r)の部分ベクトルはこの部分空間に対する検索条件（即ち部分内積がfより大きい）を満たさないため、これらの区分の検索を省略することができる。

【0071】

表Wを求めるには、部分ベクトルvが領域d内で最も部分質問ベクトルqに近い場合、p, q, vが一平面上にあり、かつvとqとのなす角ωが偏角区分cの範囲内でもっとも小さい場合を考えればよい。このとき、pとqのなす角をθ, pとvのなす角の最大値をφとすると、vとqのなす角ωは $\omega = \theta - \phi$ となるから、

$$f < v \cdot q = |v| * |q| * \cos(\theta - \phi) < R[r+1] * |q| * (\cos \theta * \cos \phi + \sin \theta \sin \phi)$$

と、

$$C[c] = \cos \phi$$

$$\cos \theta = (p \cdot q) / |p| * |q| = (p \cdot q) / |q|$$

とから、p・qが満足する不等式、

$$f < R[r+1] * C[c] * (p \cdot q) + R[r+1] * \sqrt{1 - C[c]^2} * \sqrt{|q|^2 - (p \cdot q)^2})$$

を解いて、ステップ1307のW [c, r]の式を得る。

【0072】

このように、実際の部分ベクトル v の成分を参照せず、また領域 d にも依存せずに、部分質問ベクトルのノルム $|q|$ だけから、表 $W[c, r]$ の値を定めることができる。本実施の形態の場合には、ノルム区分表 R 、偏角区分表 C がそれぞれ図 15、16 の通りなので、 $b = 0$ の場合には、表 W は図 17 のような内容となる。図中で表の値が「9. 99999」となっている要素は、この部分質問ベクトル q については、ノルムが小さすぎて、どの方向の部分ベクトル v でも、 q との内積が f に達することができないため、検索対象とはならないノルム区分であることを意味している。図 17 からは、 $c = 0$ 、即ち偏角の値が大きい場合には、広範囲な検索を行い、 $c = 3$ 、即ち偏角の値が小さい場合には、ノルムの大きな部分のみの、より狭い範囲を検索するような内容になっていることがわかる。

【0073】

ステップ 1308 では、着目する領域の中心ベクトル p と部分質問ベクトル q との内積 t を求め、偏角区分用のループ変数 c を 0 に初期化した後、ステップ 1309 で表 W の最小値を与える要素 $W[0, 255]$ より内積 t が小さいか否か検査し、小さい場合には、領域 d を鍵の一部とするどの部分ベクトルも検索条件を満足しないことが確定するのでステップ 1312 に飛び、そうでない場合にはステップ 1310 で偏角区分 c について、検索すべきノルム区分の最低値 r をステップ 1307 で算定した表 W を手がかりにして求め、この r と部分空間番号 b 、領域番号 d 、偏角区分番号 c とから、ベクトル索引 301 の検索範囲 $[k_{min}, k_{max}]$ を求める。ステップ 1311 ではこの検索範囲 $[k_{min}, k_{max}]$ を鍵にして探索木を範囲検索し、範囲検索結果に含まれる、識別番号 j とベクトル v の組 (j , v) のそれについて、部分質問ベクトル q と v との内積から部分内積下限値 f を減じた部分内積差分値を算定、内積差分表 307 の該当要素 $S[j]$ に累算する。

【0074】

例えば、 $b = 0$, $d = 4212$ の場合には、

$q = (+0.029259 -0.016005 -0.021118 +0.024992 -0.006860 -0.009032 -0.007255 -0.007715)$

$p_0 = (+1/2, -1/2, -1/2, +1/2, 0, 0, 0, 0)$

であるから、

$$t = p \cdot q = +0.045687$$

となり、 $W[0, 255] = -0.02527$ よりも t が大きく、ステップ 1310 に進む。図 17 の表 W より、

$$W[0, r] \leq t < W[0, r+1]$$

となるノルム区分番号 r は $r = 1$ であり、 $c = 0$ の場合には、探索木の鍵は、

$$\begin{aligned} [k_{\min}, k_{\max}] &= [0*6717440+4212*1024+0*256+1, 0*6717440+4212*1024+0 \\ &\quad *256+255] \\ &= [4313089, 4313343] \end{aligned}$$

となる。識別番号 1 のベクトルデータの $b = 0$ の部分ベクトル

$v = (+0.029259 -0.016005 -0.021118 +0.024992 -0.006860 -0.009032 -0.0 \\ 07255 -0.007715)$ は、鍵 $k = 0*6717440+4212*1024+0*256+1 = 4313089$ で登録さ
れているので、この範囲検索の結果の 1 つとなる。その部分内積差分値は、

$$(v \cdot q) - f = 0.221795 - 0.177436 = 0.044359$$

となり、 $S[1] = 0.044359$ となる。

また、識別番号 2 のベクトルデータの $b = 0$ の部分ベクトル

$v = (+0.029259 -0.016005 -0.021118 +0.024992 -0.006860 -0.009032 -0.0 \\ 07255 -0.007715)$ は、鍵 $k = 0*6717440+619*1024+2*256+2$ で登録されており、
 $b = 0, c = 2, d = 619$ の範囲検索の結果に含まれ、その部分内積差分値は

$$(v \cdot q) - f = 0.00005$$

となり、 $S[2] = 0.00005$ となる。

同様に、 $b = 1$ では、識別番号 2 のベクトルデータの部分ベクトルは鍵 $k = 1* \\ 6717440+2691*1024+1*256+93$ で登録されており、 $b = 1, c = 1, d = 2691$
の範囲検索の結果に含まれ、その部分内積差分値は、

$$(v \cdot q) - f = 0.00217$$

が $S[2]$ に累算され、 $S[2] = 0.00222$ となる。

【0075】

このようにして、ステップ 1312、ステップ 1313 で c を増やしながら各

偏角区分について検索範囲の決定と検索処理、内積差分の算出と累算を行う。次いで、ステップ1314、ステップ1315で領域番号dを6560まで順に増やしながら各領域についてステップ1308～ステップ1313の処理を行う。さらにステップ1316、ステップ1317で部分空間番号を37まで順に増やしながら各部分空間についてステップ1307～ステップ1315の処理を行い、類似ベクトル検索の第1ステップを終了する。この段階で、内積差分表307には各識別番号のベクトルデータVについて、質問ベクトルQとの内積 $V \cdot Q$ と、内積下限値 α との差である、内積差分値 $(V \cdot Q) - \alpha$ の推定値の上限が求まる。なぜなら、各部分空間bについて、部分質問ベクトルqとの内積が部分内積下限値fより大きな部分ベクトルについては、もれなく部分内積差分値を求めているので、部分内積差分値を求めなかったベクトルデータの部分内積差分値は負の値となるはずであり、この負の値を0で置き換えて累算している（「内積差分表を変化させない」ということは0を累算していることに等価である）ので、部分内積差分値の累算結果は、内積差分値を上から押さえる、内積差分上限値の1つとなるからである。以上のようにして内積差分表307が求めた後、類似ベクトル検索の第2ステップを実行し、最終的な検索結果を得る。

【0076】

（類似ベクトル検索の第2ステップ）

次に、図9の流れ図にそって第2ステップの処理手順を説明する。ステップ1401で現時点での検索条件を満たす候補の件数を0クリアし、ベクトルデータの内積差分を求めたか否かのフラグA[0..200000]も0、即ち「内積差分を求めていない」に初期設定する。そして、現時点で検索条件を満たす候補のうちの内積差分値の最低値（=しきい値）tを0に初期化する。

【0077】

ステップ1402では、まだ未調査、即ち内積差分を求めていないベクトルデータがあるかどうか調べ、全てのベクトルデータの内積差分を求めた場合にはステップ1412に飛ぶ。ただし、検索条件として与えた内積下限値が0以上で、かつ、ベクトルデータの各成分の分布に偏りが小さい場合には、全てのベクトルデータの内積差分を求めるはるか前にステップ1404で条件が「no」になる

ため、通常の検索条件ではステップ1402が「no」になることはない。

【0078】

ステップ1403ではA[j]が0、即ち未調査のベクトルデータ中で、内積差分表の値S[j]が最大となるようなベクトルデータの識別番号jを求める。このステップの処理は、内積差分表307を、内積差分値の降順に整列するか、あるいはヒープなどのデータ構造で表現することにより、効率のよい実行が可能である。

【0079】

ステップ1404では、先に求めたtとS[j]とを比較し、もしもS[j]がt以下であれば、未調査のベクトルデータ中には現時点のn件の候補の内積差分値を超えるベクトルデータは存在しないことが確定するのでステップ1412に飛び、現時点での候補から結果を計算して出力し、検索処理を終了する。tがS[j]より大きい場合には、ステップ1405で着目ベクトルデータのフラグA[j]を1に変えて「内積差分を求めた」ことを記録し、ベクトル索引301を検索して識別番号jのベクトルVを求め、質問ベクトルQとの内積差分値(V·Q) - α を求めて内積差分表207の該当要素S[j]内の上限値を、正確な内積差分値で置き換える。記憶領域に余裕のある場合には、内積差分表を置き換えず、新たな表に記録しても差し支えない。

【0080】

ステップ1406では、置き換えた後のS[j]とtとを再び比較し、S[j]がtより大きい場合には、ステップ1407～ステップ1411までを実行して識別番号jのベクトルデータを候補に加える。ステップ1407では、この時点で既にL件の候補が求まっているかどうか判定し、まだL件求まっていない場合には、ステップ1408で候補の件数nを増やす。ステップ1409では候補の識別番号の配列Bの最終候補(内積差分が候補中でもっとも低いような候補)としてjを登録した後、B[0..n-1]を、S[B[k]]の降順に整列する。ステップ1410で候補件数nがLに達していれば、ステップ1411でしきい値tを更新し、ステップ1402に戻って処理を続ける。

【0081】

ステップ1402、もしくはステップ1404で判定が「no」になると、上記のループを抜け、ステップ1412に進む。ステップ1412では、n件（最高L件）の候補の識別番号B[0]～B[n-1]のそれぞれについて、既に求めた内積差分値S[B[k]]に α を加えて内積の値を求め、0からn-1までの各kについて、内積がk番目に大きいベクトルデータの結果番号B[k]と、質問ベクトルVとの内積の値S[B[k]]との組（B[k], S[B[k]]）を類似ベクトル検索の最終結果として出力し、類似ベクトル検索を終了する。

【0082】

検索条件中の内積下限の値が0.5以上と十分に大きく、かつベクトルデータの分布に大きな偏りがなく、内積下限 α 以上の内積となるベクトルデータの個数が取得件数Lよりも十分多い場合には、ステップ1402～ステップ1411のループを、取得件数Lの数倍程度繰り返すと、ステップ1404の判定が「no」となるため、実際にベクトルを検索して内積を求めるベクトルデータの個数が非常に少なく、最終結果を効率よく求めることが可能となる。しかも、この特性はしが数百程度の場合でも成り立つため、Lが比較的大きい検索条件の場合には、Lが高々数個の場合にしか実用的な検索速度が得られない従来の類似ベクトル検索方法に比べ、処理効率が非常によくなる。

【0083】

以上説明したように、本発明の第3の実施の形態の類似ベクトル検索方法および装置によれば、ベクトルの次元が数百に及ぶベクトルデータを多数集めたベクトルデータベースに対して、「最も類似しているL個のベクトルデータを求める」という型の高速な類似検索が可能で、さらにしが比較的大きい（数十～数百個）場合でも検索処理が極端に遅くならず、「内積の値が0.8以上」といった類似検索範囲の指定が可能な、ベクトルの内積を類似尺度として用いる優れた類似ベクトル検索方法および装置を提供することができる。

【0084】

なお、本実施の形態3においては、本発明の実施の形態1のベクトル索引作成装置で作成したベクトル索引を検索する場合について説明したが、実施の形態1の索引作成装置において、各部分ベクトルを求める処理を、ノルム区分番号と各

成分の成分区分番号とから各成分の値を求めるように変更するだけで、本実施の形態3の類似ベクトル検索装置は、実施の形態2のベクトル索引作成装置で作成したベクトル索引の検索にも用いることができ、上記と同様な効果が期待できる。

【0085】

さらに、本実施の形態3においては、類似ベクトル検索の第1ステップにおいて、各部分空間 b についての検索処理を逐次的に行う手順を説明したが、図8の流れ図のステップ1306～ステップ1317のループは、多数のCPU（中央処理装置）を持つ並列計算機の場合には、それぞれのCPUで分担して処理し、共通の内積差分表に中間結果として累算することで、容易に高い並列度で並列処理することができ、検索速度をさらに向上させることができある。

【0086】

(実施の形態4)

次に、本発明の第4の形態について、図面を参照しながら説明する。

(類似ベクトル検索装置の構成)

図4は本発明の請求項10、11、13、23、24、26にかかる、類似ベクトル検索装置の全体構成を表わすブロック図である。図4において、ベクトル索引401は、上記した第1の実施の形態のベクトル索引作成装置によって作成したものであり、20万件の新聞記事を集めた新聞記事全文データベースから作成した、各新聞記事の特徴を表す296次元の実ベクトルと、各記事を一意に識別するための1から200000の範囲の識別番号の2項目からなるベクトルデータを20万個格納した、図12のような内容を持つベクトルデータベースから作成したベクトル索引である。

【0087】

検索条件入力手段402は、前記新聞記事全文データベースに対する類似検索を行うために、前記新聞記事全文データベース中のいずれかの記事の識別番号と、類似検索範囲を示す0～100までの類似度下限値と最大取得件数とを入力し、ベクトル索引401を識別番号で検索して入力された識別番号から対応する記事のベクトルを質問ベクトルQとして求め、類似度下限値から二乗距離、即ち距離

の二乗の上限値である、二乗距離上限値 α を求める。

【0088】

部分質問条件算定手段403は、検索条件入力手段402が求めた質問ベクトルQに対し、0～36までの部分空間について、37種の8次元部分質問ベクトルqと、qに対応する部分ベクトルとの二乗距離の上限値である部分二乗距離上限値fを $f = \alpha |q|^2 / |Q|^2$ によって算定する。

【0089】

検索対象範囲生成手段404は、部分質問条件算定手段403が求めた部分空間bに対する部分質問ベクトルqと部分二乗距離上限値fとベクトル索引401中のノルム区分表および偏角区分表とから、部分質問ベクトルqとの部分二乗距離が部分二乗距離上限値fより小さい可能性がある部分ベクトルを含んでいる領域を特定する領域番号d、偏角区分番号c、ノルム区分範囲 $[r_1, r_2]$ の組(d, c, $[r_1, r_2]$)を全て列挙する。

【0090】

索引検索手段405は、検索対象領域生成手段404が各部分空間bに対して生成した(d, c, $[r_1, r_2]$)から、ベクトル索引401に対する検索条件Kを、ベクトル索引作成時における鍵の算定と同様に、

$$K = [k_{\min}, k_{\max}]$$

$$k_{\min} = b * 7617440 + d * 1024 + c * 256 + r_1$$

$$k_{\max} = b * 7617440 + d * 1024 + c * 256 + r_2$$

と算定し、ベクトル索引401を検索条件Kで範囲検索して検索条件に合致する鍵を持つ部分ベクトルvと識別番号iとの組(i, v)をすべて求める。

【0091】

二乗距離差分上限算定手段406は、索引検索手段405が求めた部分ベクトルvと識別番号iとの組(i, v)および部分質問条件算定手段403が求めた部分質問ベクトルqと部分二乗距離上限値fから、 $t = f - |v - q|^2$ によって部分二乗距離差分値tを算定して、識別番号iを添字とする表の要素S[i]に、部分二乗距離差分値tを累算(加算)することで、二乗距離上限値 α から、識別番号iのベクトルデータのベクトルVと質問ベクトルQとの二乗距離 $|V -$

$|Q|^2$ を減じた二乗距離差分の上限値を算定する。

【0092】

二乗距離差分表407は、二乗距離差分上限算定手段406が二乗距離差分の上限値を蓄積する、識別番号*i*のベクトルベータの二乗距離差分値S[i]を参照／格納する。

【0093】

類似結果決定手段408は、二乗距離差分表407の要素S[i]中で、二乗距離差分上限値S[i]が正の値で大きいものから順に、ベクトル索引401を識別番号*i*で検索して対応するベクトルVを求め、検索条件入力手段402が算出した二乗距離上限値 α から、Vと検索条件入力手段402が算出した質問ベクトルQとの二乗距離 $|V - Q|^2$ を差し引いた二乗距離差分値 $\alpha - |V - Q|^2$ を算出して、S[i]を二乗距離差分値 $\alpha - |V - Q|^2$ で置き換え、二乗距離差分値を算出していない記事の部分二乗距離差分累算値の最大値より大きい二乗距離差分値を持った二乗距離差分値を算出した記事がL個以上になった時点、あるいは部分二乗距離差分累算値が正である全ての記事の二乗距離差分値を算出した時点で、二乗距離差分値が正である、二乗距離差分値の大きな最大L個の結果候補(i, S[i])に対し、識別番号*i*と距離sqrt($\alpha - S[i]$)との組(i, sqrt($\alpha - S[i]$))を、検索結果として、検索結果出力手段に出力する。

【0094】

検索結果出力手段409は、類似ベクトル検索の結果である、最大L件の新聞記事の識別番号と0~100の範囲の類似度とを、類似結果決定手段408が求めた検索結果から算定して表示するものである。

【0095】

(類似ベクトル検索装置の動作)

以上の構成の類似ベクトル検索装置において、その動作を図面をもとに説明する。図10は類似ベクトル検索の第1ステップにおける検索処理の手順を記述した流れ図、図11は類似ベクトル検索の第2ステップにおける検索処理の手順を記述した流れ図である。類似ベクトル検索の第1ステップでは、検索条件入力手段402から入力された検索条件から部分質問ベクトルqと部分二乗距離上限値

f とを作成し、ベクトル索引401を検索して、各ベクトルデータの二乗距離差分上限値 $S[i]$ 、即ち二乗距離上限値から質問ベクトルとの二乗距離を減じた値が $S[i]$ 未満となるような値を二乗距離差分表407に求める。次に、類似ベクトル検索の第2ステップでは、第1ステップで二乗距離差分表407に求めた二乗距離差分上限値を手がかりにして、類似検索結果決定手段408が、「質問ベクトルとの二乗距離が α より小さい」という検索条件に合致し、かつ質問ベクトルとの二乗距離が比較的小さいベクトルデータから順に、ベクトルの成分を検索して二乗距離差分を求めていき、二乗距離差分を求めていないベクトルデータのどれよりも二乗距離差分値が大きいことが保証されるベクトルデータが指定件数（即ち L 件）以上集まるか、あるいは検索条件に合致する全てのベクトルデータについて二乗距離差分値を求めるまで処理を続け、得られた二乗距離差分値から距離を算出して最終結果を出力する。

【0096】

(類似ベクトル検索の第1ステップ)

以下、検索条件として識別番号1、類似度下限値90、最大取得件数10が入力された場合を例にとって、図10、図11にそって類似ベクトル検索の内容を説明する。識別番号が1なので、296次元のベクトルの各成分は、図12の上半分の通りとなる。まず、ステップ1501で二乗距離差分表 S の20万個の要素 $S[0] \sim S[200000]$ を0に初期設定する。次に、上記の検索条件を検索条件入力手段402から読み込み、それぞれ i , Z , L に格納する。

【0097】

ステップ1502で部分空間番号 b を0に初期化したあと、類似度下限値 Z から、二乗距離上限値 α を算出する。この検索条件の場合には、 $\alpha \leftarrow (100 - 90) / 50 = 0.2$ となる。ステップ1504、ステップ1505で、各部分空間について、ベクトル索引401の逆変換表 K を引いて鍵を求め、探索表を検索してベクトルデータを求め、識別番号が1であるもののベクトル部分を Q に格納することで、質問ベクトルを $Q[0..295]$ に求める。ステップ1506で部分空間番号を初期化した後、ステップ1507～ステップ1517で各部分空間についてベクトル索引を検索して二乗距離差分表407に各ベクトルデータの二乗距

離差分上限値を求める。

【0098】

ステップ1507では、部分空間番号 b の部分質問ベクトル q [0..7] と部分二乗距離上限値 f 、即ちこの部分空間の部分ベクトルデータ v と q との部分二乗距離の上限値を求める。 $b = 0$ の場合には、 $|q|^2 = 0.221795$, $|Q|^2 = 1$ であるから、

$$f = 0.2 * 0.221795 / 1.0 = 0.044359$$

となる。領域番号 d を 0 に初期化したあと、検索対象範囲を決定するために用いる表 W を作成する。表 W は偏角区分番号 c 、ノルム区分番号 r で引いて、着目している領域番号 d の領域の中心ベクトル p と部分質問ベクトル q との内積 $p \cdot q$ が $W[c, r]$ 未満ならば、区分 $(d, c, 0) \sim (d, c, r)$ の部分ベクトル v と部分質問ベクトル q との部分二乗距離が f 以上となるように作成する。この場合には、区分 $(d, c, 0) \sim (d, c, r)$ の部分ベクトルはこの部分空間に対する検索条件（即ち部分二乗距離が f より小さい）を満たさないため、これらの区分の検索を省略することができる。

【0099】

表 W を求めるには、部分ベクトル v が領域 d 内で最も部分質問ベクトル q に近い場合、即ち p, q, v が一平面上にあり、かつ v と q とのなす角 ω が偏角区分 c の範囲内でもっとも小さい場合を考えればよい。このとき、 p と q のなす角を θ 、 p と v のなす角の最大値を ϕ とすると、 v と q のなす角 ω は $\omega = \theta - \phi$ となるから、

$$\begin{aligned} f > |v - q|^2 &= |v|^2 + |q|^2 - 2 * |v| * |q| * \cos(\theta - \phi) \\ &> R[r]^2 + |q|^2 - 2 * R[r+1] * |q| * (\cos \theta * \cos \phi + \\ &\quad \sin \theta \sin \phi) \end{aligned}$$

と、

$$C[c] = \cos \phi$$

$$\cos \theta = (p \cdot q) / |p| * |q| = (p \cdot q) / |q|$$

とから、 $p \cdot q$ が満足する不等式

$$f > R[r]^2 + |q|^2 - 2 * R[r+1] * ((p \cdot q) * C[c] + \sqrt{|q|^2 - (p \cdot q)^2}) *$$

$$\sqrt{1 - C [c]^2})$$

を解いて、ステップ1507のW [c, r] の式を得る。

【0100】

このように、実際の部分ベクトルvの成分を参照せず、また領域dにも依存せずに、部分質問ベクトルのノルム |q| だけから表W [c,r] の値を定めることができる。本実施例の場合には、ノルム区分表R、偏角区分表Cがそれぞれ図15、図16の通りなので、b = 0, b = 1の場合には、表Wは図18のような内容となる。図17と同様に、図中で表の値が「9. 99999」となっている要素は、この部分質問ベクトルqについては検索対象とはならないノルム区分であることを意味している。また、b = 0では区分10～255, b = 1区分0～59および区分180～255までの表の値が記載されていないが、この部分はすべて値が「9. 99999」であるので省略している。これが、類似尺度として距離を用いているため、ノルムが小さすぎても、逆に大きすぎても、部分質問ベクトルとの距離が大きくなってしまい、結果的に「距離がα未満」という検索条件を満足できなくなるためである。

【0101】

ステップ1508で着目する領域の領域中心ベクトルpと部分質問ベクトルqとの内積tを求め、偏角区分用のループ変数cを0に初期化した後、ステップ1509で表Wの最小値を与える要素Min (W [0, r]) より内積tが小さいか否か検査し、小さい場合には、領域dを鍵の一部とするどの部分ベクトルも検索条件を満足しないことが確定するのでステップ1512に飛び、そうでない場合には、ステップ1510で偏角区分cについて、検索すべきノルム区分の最低値r_min及び最高値r_maxをステップ1507で算定した表Wを手がかりにして、W [c, r] < t が成り立つノルム区分番号rの区間として求め、この[r_min, r_max]と部分空間番号b、領域番号d、偏角区分番号cとから、ベクトル索引401の検索範囲[k_min, k_max]を求める。

【0102】

ステップ1511では、この検索範囲 [k_min, k_max] を鍵にして探索木を範囲検索し、範囲検索結果に含まれる、識別番号jとベクトルvの組 (j, v) の

それぞれについて、部分二乗距離上限値 f から部分質問ベクトル q と v の部分二乗距離 $|v - q|^2$ を減じた部分二乗距離差分値を算定、二乗距離差分表407の該当要素 $S[j]$ に累算する。例えば、 $b = 0$, $d = 4212$ の場合には、

$$q = (+0.029259 -0.016005 -0.021118 +0.024992 -0.006860 -0.009032 -0.007255 -0.007715)$$

$$p = (+1/2, -1/2, -1/2, +1/2, 0, 0, 0, 0)$$

であるから、

$$t = p \cdot q = +0.045687$$

となり、 $\text{Min}(W[0, r]) = 0.03356$ よりも t が大きく、ステップ1510に進む。図15の表Wより、例えば $c = 0$ の場合には、

$$r_{\min} = 1, \quad r_{\max} = 5$$

であり、探索木の検索範囲は、

$$[k_{\min}, k_{\max}] = [0*6717440+4212*1024+0*256+1, 0*6717440+4212*1024+0*256+5] = [4313089, 4313093]$$

となる。識別番号1のベクトルデータの $b = 0$ の部分ベクトル x は、

$$x = (+0.029259 -0.016005 -0.021118 +0.024992 -0.006860 -0.009032 -0.007255 -0.007715)$$

となり、 $k = 0*6717440+4212*1024+0*256+1 = 4313089$ で登録されているので、この範囲検索の結果の1つとなる。その部分二乗距離差分値は、

$$f - |v - q|^2 = 0.044359 - 0 = 0.044359$$

となり、 $S[1] = 0.044359$ となる。

--【0103】--

このようにして、ステップ1512、ステップ1513で c を増やしながら各偏角区分について検索範囲の決定と検索処理、二乗距離差分の算出と累算を行う。ステップ1514、ステップ1515で領域番号 d を6560まで順に増やしながら各領域についてステップ1508～ステップ1513の処理を行う。ステップ1516、ステップ1517で部分空間番号を37まで順に増やしながら各部分空間についてステップ1507～ステップ1515の処理を行い、類似ベクトル検索の第1ステップを終了する。この段階で、二乗距離差分表507には各

識別番号のベクトルデータVについて、二乗距離上限値 α と、質問ベクトルQとの二乗距離 $|V-Q|^2$ との差である、二乗距離差分値 $\alpha - |V-Q|^2$ の推定値の上限が求まる。なぜなら、各部分空間bについて、部分質問ベクトルqとの二乗距離が部分二乗距離上限値fより小さな部分ベクトルについては、もれなく部分二乗距離差分値を求めており、部分二乗距離差分値を求めなかったベクトルデータの部分二乗距離差分値は負の値となるはずであり、この負の値を0で置き換えて累算している（「二乗距離差分表を変化させない」ということは0を累算していることに等価である）ので、部分二乗距離差分値の累算結果は、二乗距離差分値を上から押さえる、二乗距離差分上限値の1つとなるからである。こうして二乗距離差分表407が求めた後、類似ベクトル検索の第2ステップを実行し、最終的な検索結果を得る。

【0104】

（類似ベクトル検索の第2ステップ）

次に、図11の流れ図にそって第2ステップの処理手順を説明する。ステップ1601で現時点での検索条件を満たす候補の件数を0クリアし、ベクトルデータの二乗距離差分を求めたか否かのフラグA[0..200000]も0、即ち「二乗距離差分を求めていない」に初期設定する。そして、現時点で検索条件を満たす候補のうちの二乗距離差分値の最低値（=しきい値）tを0に初期化する。

【0105】

ステップ1602では、まだ未調査、即ち二乗距離差分を求めていないベクトルデータがあるかどうか調べ、全てのベクトルデータの二乗距離差分を求めた場合にはステップ1612に飛ぶ。ただし、検索条件として与えた二乗距離上限値が1以下で、かつ、ベクトルデータの各成分の分布に偏りが小さい場合には、全てのベクトルデータの二乗距離差分を求めるはるか前にステップ1604で条件が「no」になるため、通常の検索条件ではステップ1602が「no」になることはない。ステップ1603ではA[j]が0、即ち未調査のベクトルデータ中で、二乗距離差分表の値S[j]が最大となるようなベクトルデータの識別番号jを求める。このステップの処理は、内積差分表407を、二乗距離差分値の降順に整列するか、あるいはヒープなどのデータ構造で表現することにより、効

率のよい実行が可能である。

【0106】

ステップ1604では、先に求めた t と $S[j]$ とを比較し、もしも $S[j]$ が t 以下であれば、未調査のベクトルデータ中には現時点の n 件の候補の二乗距離差分値を超えるベクトルデータは存在しないことが確定するのでステップ1612に飛び、現時点での候補から結果を計算して出力し、検索処理を終了する。 t が $S[j]$ より大きい場合には、ステップ1605で着目ベクトルデータのフラグ $A[j]$ を1に変えて「二乗距離差分を求めた」ことを記録し、ベクトル索引401を検索して識別番号 j のベクトル V を求め、質問ベクトル Q との二乗距離差分値 $\alpha - |V - Q|^2$ を求めて二乗距離差分表407の該当要素 $S[j]$ 内の上限値を、正確な二乗距離差分値で置き換える。記憶領域に余裕のある場合には、二乗距離差分表を置き換えず、新たな表に記録しても差し支えない。ステップ1606で、置き換えた後の $S[j]$ と t とを再び比較し、 $S[j]$ が t より大きい場合には、ステップ1607～ステップ1611までを実行して識別番号 j のベクトルデータを候補に加える。

【0107】

ステップ1607では、この時点で既に L 件の候補が求まっているかどうか判定し、まだ L 件求まっていない場合には、ステップ1608で候補の件数 n を増やす。ステップ1609では候補の識別番号の配列 B の最終候補（二乗距離差分が候補中でもっとも低いような候補）として j を登録した後、 $B[0..n-1]$ を、 $S[B[k]]$ の降順に整列する。ステップ1610で候補件数 n が L に達していれば、ステップ1611でしきい値 t を更新し、ステップ1602に戻って処理を続ける。ステップ1602、もしくはステップ1604で判定が「no」になると、上記のループを抜け、ステップ1612に進む。

【0108】

ステップ1612では、 n 件（最高 L 件）の候補の識別番号 $B[0] \sim B[n-1]$ のそれぞれについて、既に求めた二乗距離差分値 $S[B[k]]$ から、 $sqr t(\alpha - S[B[k]])$ で質問ベクトル Q との距離を求め、0から $n-1$ までの各 k について、距離が k 番目に小さいベクトルデータの結果番号 $B[k]$ と、質問ベクトル Q

との距離の値 $S[B[k]]$ との組 $(B[k], S[B[k]])$ を類似ベクトル検索の最終結果として出力し、類似ベクトル検索を終了する。

【0109】

検索条件中の二乗距離上限の値が 0.5 以下と十分に小さく、かつベクトルデータの分布に大きな偏りがなく、二乗距離上限 α 未満の二乗距離となるベクトルデータの個数が取得件数 L よりも十分多い場合には、ステップ 1602～ステップ 1611 のループを、取得件数 L の数倍程度繰り返すと、ステップ 1604 の判定が「no」となるため、実際にベクトルを検索して二乗距離を求めるベクトルデータの個数が非常に少なく、最終結果を効率よく求めることが可能となる。しかも、この特性は L が数百程度の場合でも成り立つため、 L が比較的大きい検索条件の場合には、 L が高々数個の場合にしか実用的な検索速度が得られない従来の類似ベクトル検索方法に比べ、処理効率が非常によくなる。

【0110】

以上、説明したように、本発明の第 4 の実施の形態の類似ベクトル検索方法によれば、ベクトルの次元が数百に及ぶベクトルデータを多数集めたベクトルデータベースに対して、「最も類似している L 個のベクトルデータを求める」という型の高速な類似検索が可能で、さらに L が比較的大きい（数十～数百個）場合でも検索処理が極端に遅くならず、「距離の値が 0.2 以下」といった類似検索範囲の指定が可能な、ベクトル間の距離を類似尺度として用いる優れた類似ベクトル検索方法が実現できる。

【0111】

なお、本実施の形態 4 においては、本発明の実施の形態 1 のベクトル索引作成装置で作成したベクトル索引を検索する場合について説明したが、実施の形態 1 の索引作成装置において、各部分ベクトルを求める処理を、ノルム区分番号と各成分の成分区分番号とから各成分の値を求めるように変更するだけで、本実施の形態 4 の類似ベクトル検索装置は、実施の形態 2 のベクトル索引作成装置で作成したベクトル索引の検索にも用いることができ、上記と同様な効果が期待できる。

【0112】

また、本実施の形態4においては、質問ベクトルを直接入力せず、ベクトルデータベース中のベクトルデータの識別番号を指定する形態を説明したが、外部から直接質問ベクトルのデータを指定した場合でも、上記の説明と同様の方法で、類似ベクトル検索装置を用意に実施できる。

【0113】

さらに、本実施の形態4においては、類似ベクトル検索の第1ステップにおいて、各部分空間 b についての検索処理を逐次的に行う手順を説明したが、図10の流れ図のステップ1506～ステップ1517のループは、多数のCPU（中央処理装置）を持つ並列計算機の場合には、それぞれのCPUで分担して処理し、共通の内積差分表に中間結果として累算することで、容易に高い並列度で並列処理することができ、検索速度をさらに向上させることができある。

【0114】

【発明の効果】

以上のように本発明のベクトル索引作成方法は、部分ベクトル算定手段と、ノルム分布集計手段と、ノルム区分表と、領域番号算定手段と、偏角分布集計手段と、偏角区分表と、ノルム区分番号算定手段と、偏角区分番号算定手段と、索引データ算定手段と、索引構成手段とを有することにより、方向、ノルムの分布が不明なベクトルデータベースに対して、ベクトルの次元が数百次元の場合でも、高速な検索が可能で、類似検索の際に、ベクトル間の距離と、ベクトルの内積の2種類の類似度のいずれかを選択でき、「もっとも類似しているL個のベクトルを求める」という型の類似検索ができ、さらに、Lが比較的大きい（数十～数百）場合でも、検索処理が極端に遅くならず、「内積が0.6以上」といった、類似検索範囲の指定が可能で、索引作成に要する計算量が実用的範囲であるような、ベクトル索引の作成を可能にする、という効果を奏するものである。

【0115】

また、本発明のベクトル索引作成方法は、成分区分番号算定手段をさらに有することにより、上記の効果に加えて、成分の量子化による計算誤差を最小限に抑えながら、作成すべきベクトル索引の容量を大幅に低減できる、という効果を奏するものである。

【0116】

また、本発明の類似ベクトル検索方法は、部分質問条件算定手段と、検索対象範囲生成手段と、索引検索手段と、類似候補集合算定手段と、類似検索結果決定手段とを有し、部分内積差分の累算値を算定して類似検索の手がかりとして用いることにより、ベクトルデータベースに対して、ベクトルの次元が数百次元の場合でも、高速な検索が可能で、「もっとも類似しているL個のベクトルを求める」という型の類似検索ができ、さらに、Lが比較的大きい（数十～数百）場合でも、検索処理が極端に遅くならず、「内積が0.6以上」といった、類似検索範囲の指定が可能な、内積あるいは距離を類似尺度として用いる類似ベクトル検索を可能にする、という効果を奏するものである。しかも、内積と距離のどちらを類似尺度として用いるかをベクトル索引の作成時に指定する必要がなく、単一のベクトル索引を使って、検索時に必要に応じて、類似尺度の使い分けができるという、優れた効果を奏するものである。

【0117】

また、本発明の類似ベクトル検索方法は、部分質問条件を算定する手段と、検索対象範囲を生成する手段と、索引を検索する手段と、二乗距離差分上限を算定する手段と、類似検索結果を決定する手段とを有し、部分二乗距離差分の累算値を算定して類似検索の手がかりとして用いることにより、ベクトルデータベースに対して、ベクトルの次元が数百次元の場合でも、高速な検索が可能で、「もっとも類似しているL個のベクトルを求める」という型の類似検索ができ、さらに、Lが比較的大きい（数十～数百）場合でも、検索処理が極端に遅くならず、「距離が0.8以下」といった、類似検索範囲の指定が可能で、距離を類似尺度として用いる類似ベクトル検索を可能にする、という効果を奏するものである。

【0118】

本発明の効果は、索引作成対象、検索対象となるベクトルデータの次元が数百次元と高く、ベクトルデータベース中のベクトルデータ数が数十万～数百万個と多く、検索時の取得件数が数十個と多い場合には特に顕著であり、従来のベクトル索引作成方法では数百時間をしていた索引作成時間を数十分に短縮するとともに、従来の類似ベクトル検索方法では数分を要する、あるいは実行不可能であ

った類似検索処理を1秒以下で処理することができ、実用上、非常に大きな効果が得られる。

【図面の簡単な説明】

【図1】

第1の実施の形態におけるベクトル索引作成装置の全体構成を示すブロック図

【図2】

第2の実施の形態におけるベクトル索引作成装置の全体構成を示すブロック図

【図3】

第3の実施の形態における類似ベクトル検索装置の全体構成を示すブロック図

【図4】

第4の実施の形態における類似ベクトル検索装置の全体構成を示すブロック図

【図5】

第1、第2の実施の形態におけるベクトル索引作成の第1ステップの作成手順を記述した流れ図

【図6】

第1の実施の形態におけるベクトル索引作成の第2、第3ステップの作成手順を記述した流れ図

【図7】

第2の実施の形態におけるベクトル索引作成の第2、第3ステップの作成手順を記述した流れ図

【図8】

第3の実施の形態における類似ベクトル検索の第1ステップの検索手順を記述した流れ図

【図9】

第3の実施の形態における類似ベクトル検索の第2ステップの検索手順を記述した流れ図

【図10】

第4の実施の形態における類似ベクトル検索の第1ステップの検索手順を記述した流れ図

【図11】

第4の実施の形態における類似ベクトル検索の第2ステップの検索手順を記述した流れ図

【図12】

第1、第2、第3、第4の実施の形態におけるベクトルデータベースの内容例を示す一覧図

【図13】

第1、第2の実施の形態におけるノルム分布集計結果例を示す特性図

【図14】

第1、第2の実施の形態における偏角分布集計結果例を示す特性図

【図15】

第1、第2、第3、第4の実施の形態におけるノルム区分表の内容例を示す一覧図

【図16】

第1、第2、第3、第4の実施の形態における偏角区分表の内容例を示す一覧図

【図17】

第3の実施の形態における表Wの内容例（一部分）を示す一覧図

【図18】

第4の実施の形態における表Wの内容例（一部分）を示す一覧図

【符号の説明】

101 ベクトルデータベース

102 部分ベクトル算定手段

103 ノルム分布集計手段

104 ノルム区分表

105 領域番号算定手段

106 偏角分布集計手段

107 偏角区分表

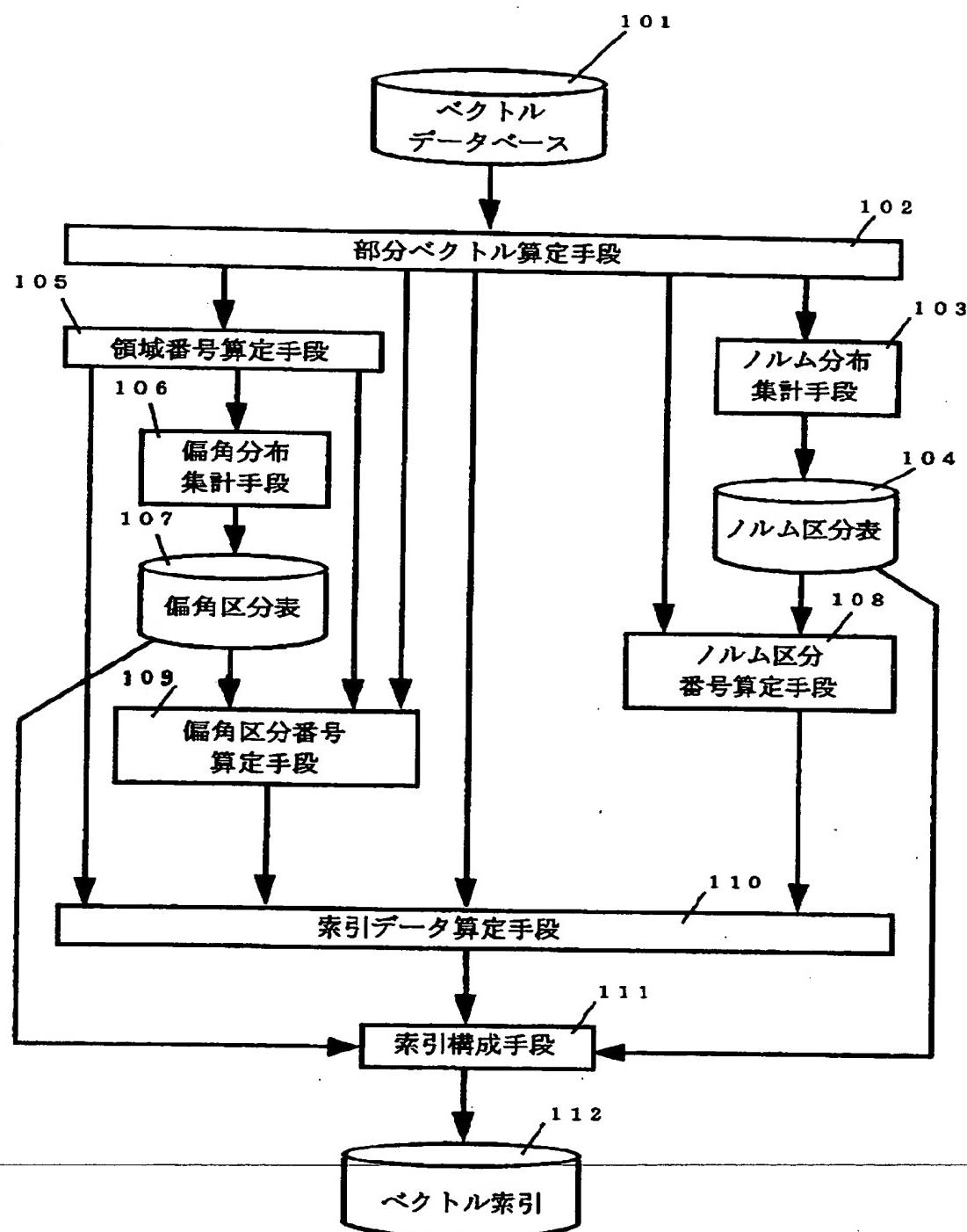
108 ノルム区分番号算定手段

- 109 偏角区分番号算定手段
- 110 索引データ算定手段
- 111 索引構成手段
- 112 ベクトル索引
- 201 ベクトルデータベース
- 202 部分ベクトル算定手段
- 203 ノルム分布集計手段
- 204 ノルム区分表
- 205 領域番号算定手段
- 206 偏角分布集計手段
- 207 偏角区分表
- 208 ノルム区分番号算定手段
- 209 偏角区分番号算定手段
- 210 索引データ算定手段
- 211 索引構成手段
- 212 ベクトル索引
- 213 成分区分番号算定手段
- 301 ベクトル索引
- 302 検索条件入力手段
- 303 部分質問条件算定手段
- 304 検索対象範囲生成手段
- 305 索引検索手段
- 306 内積差分上限算定手段
- 307 内積差分表
- 308 類似検索結果決定手段
- 309 検索結果出力手段
- 401 ベクトル索引
- 402 検索条件入力手段
- 403 部分質問条件算定手段

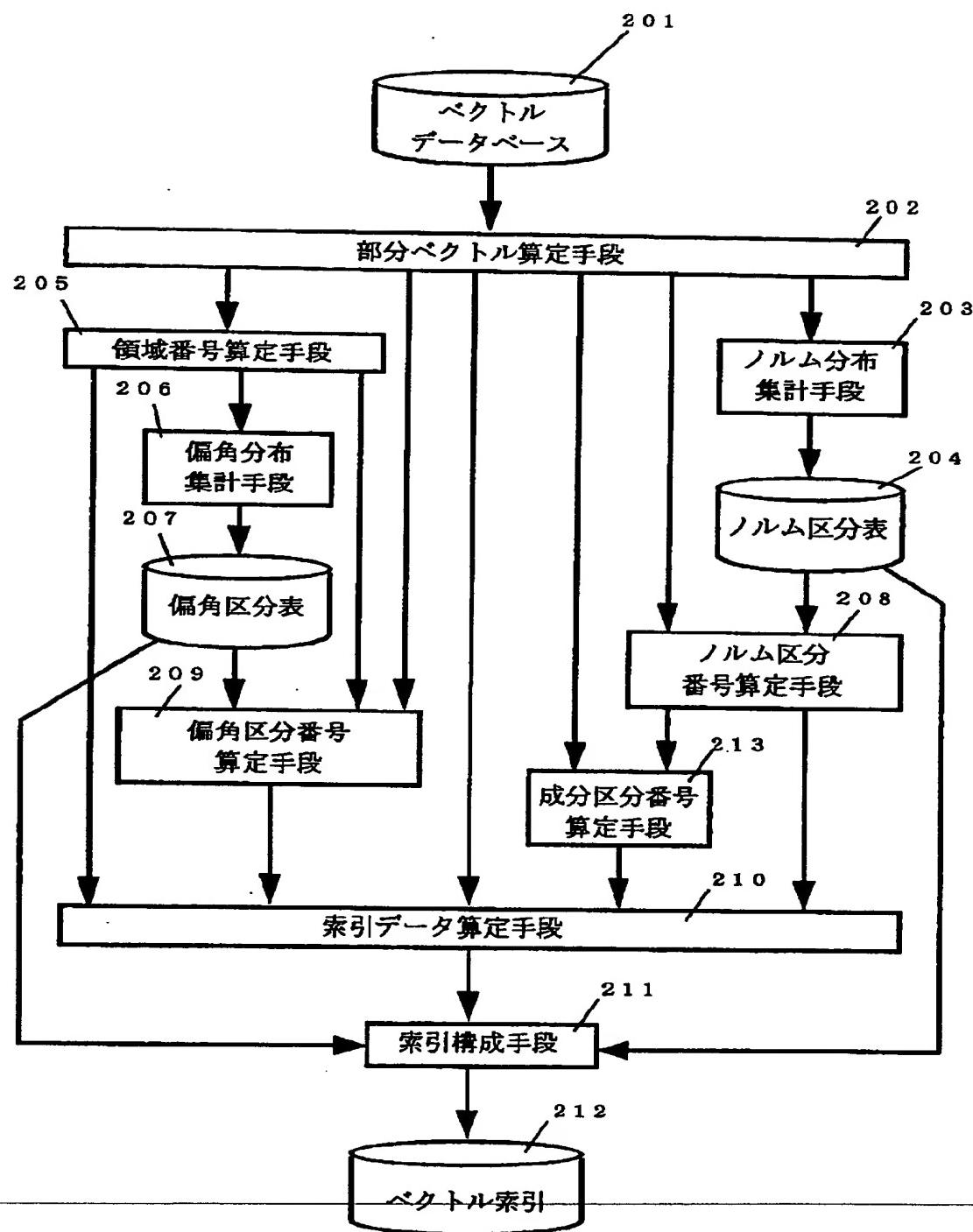
- 404 検索対象範囲生成手段
- 405 索引検索手段
- 406 二乗距離差分上限算定手段
- 407 二乗距離差分表
- 408 類似検索結果決定手段
- 409 検索結果出力

【書類名】 図面

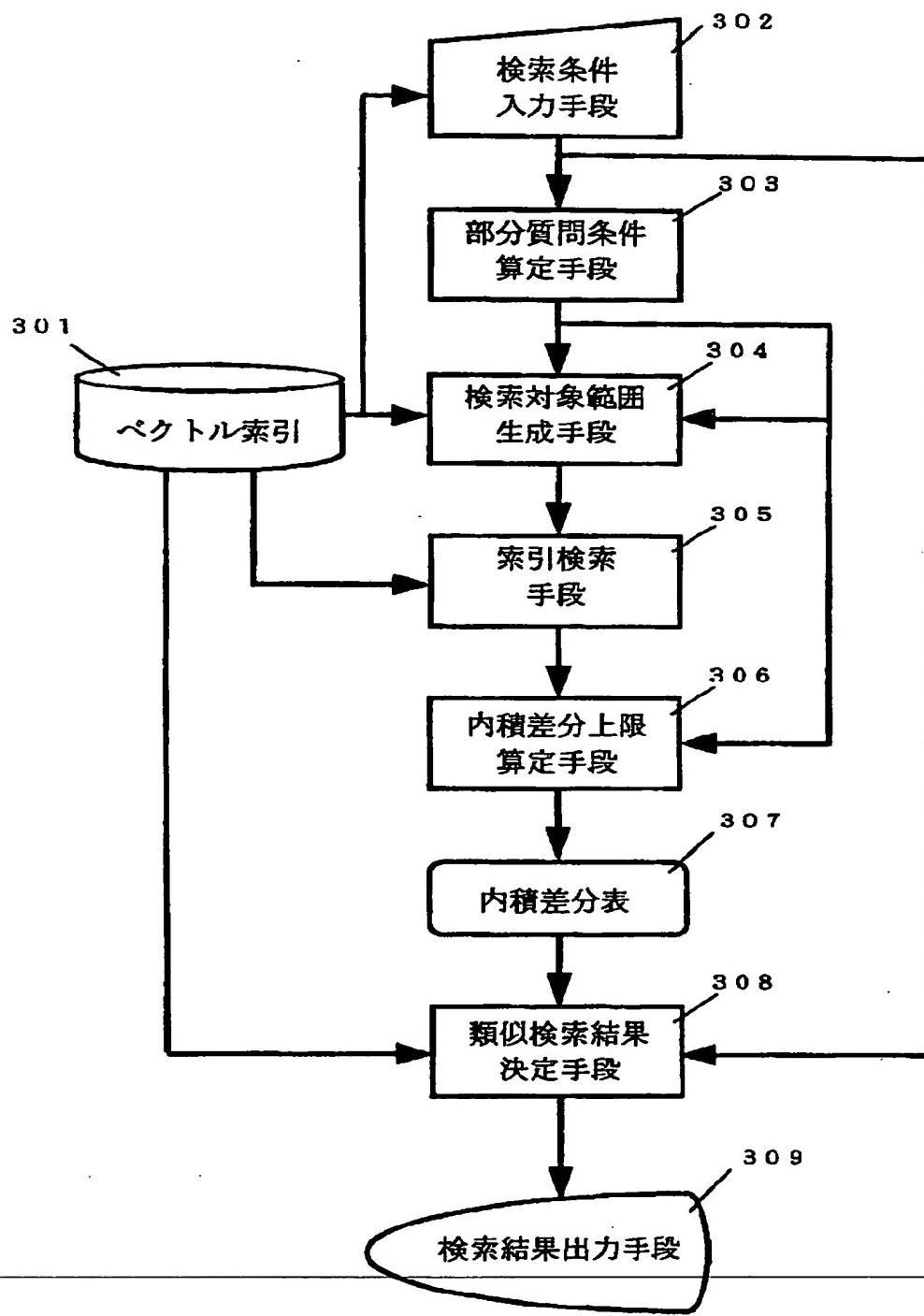
【図1】



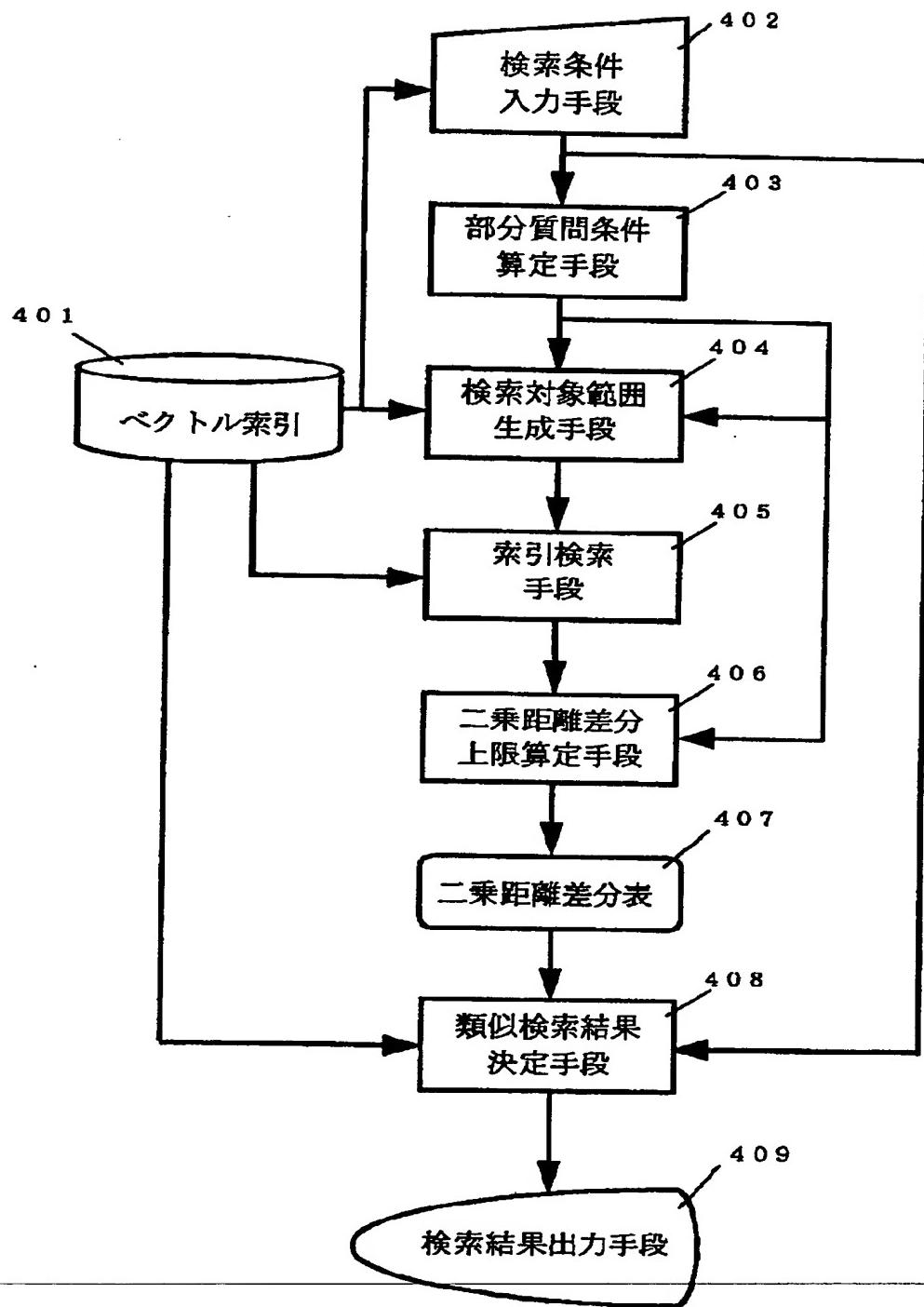
【図2】



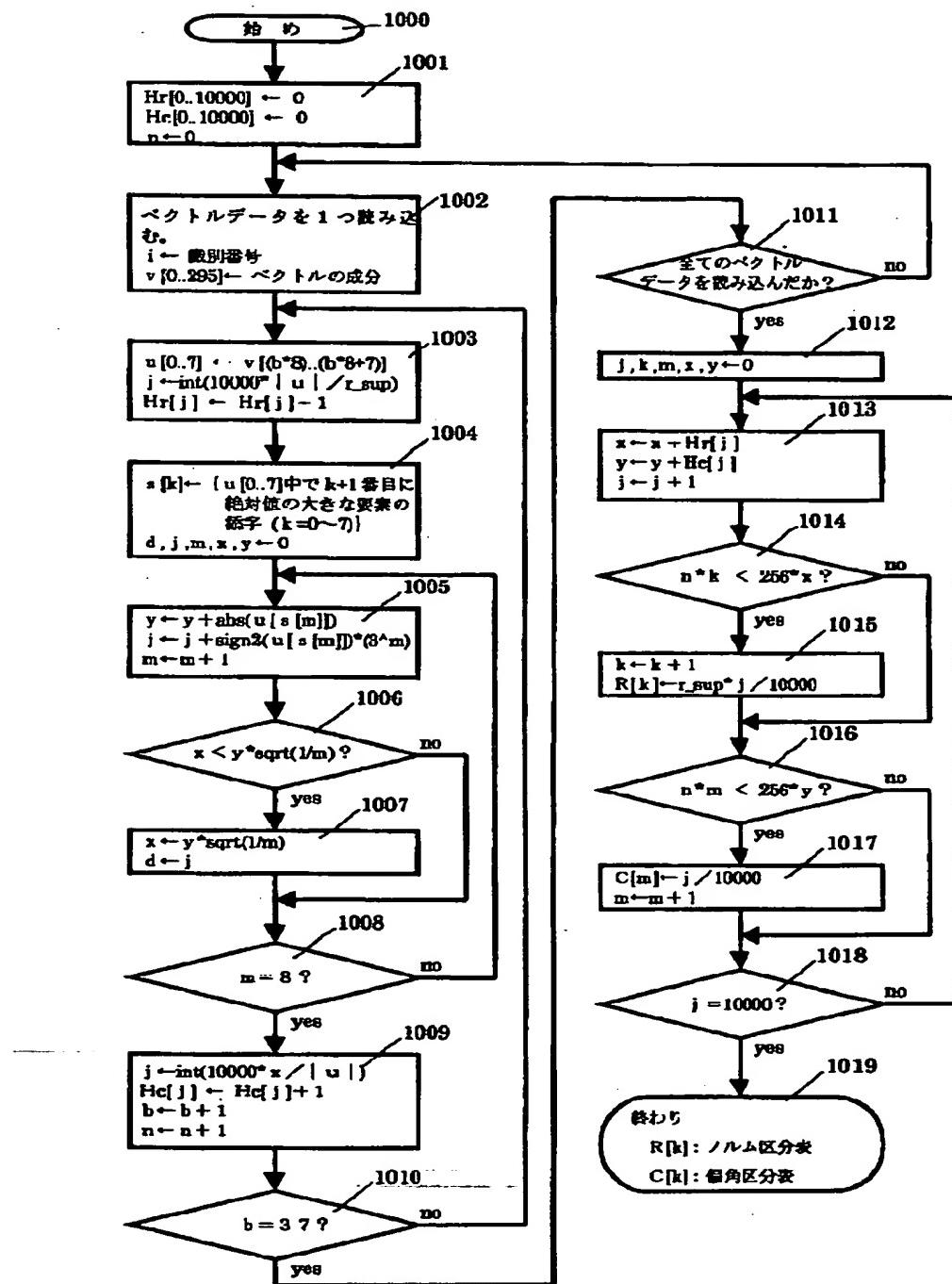
【図3】



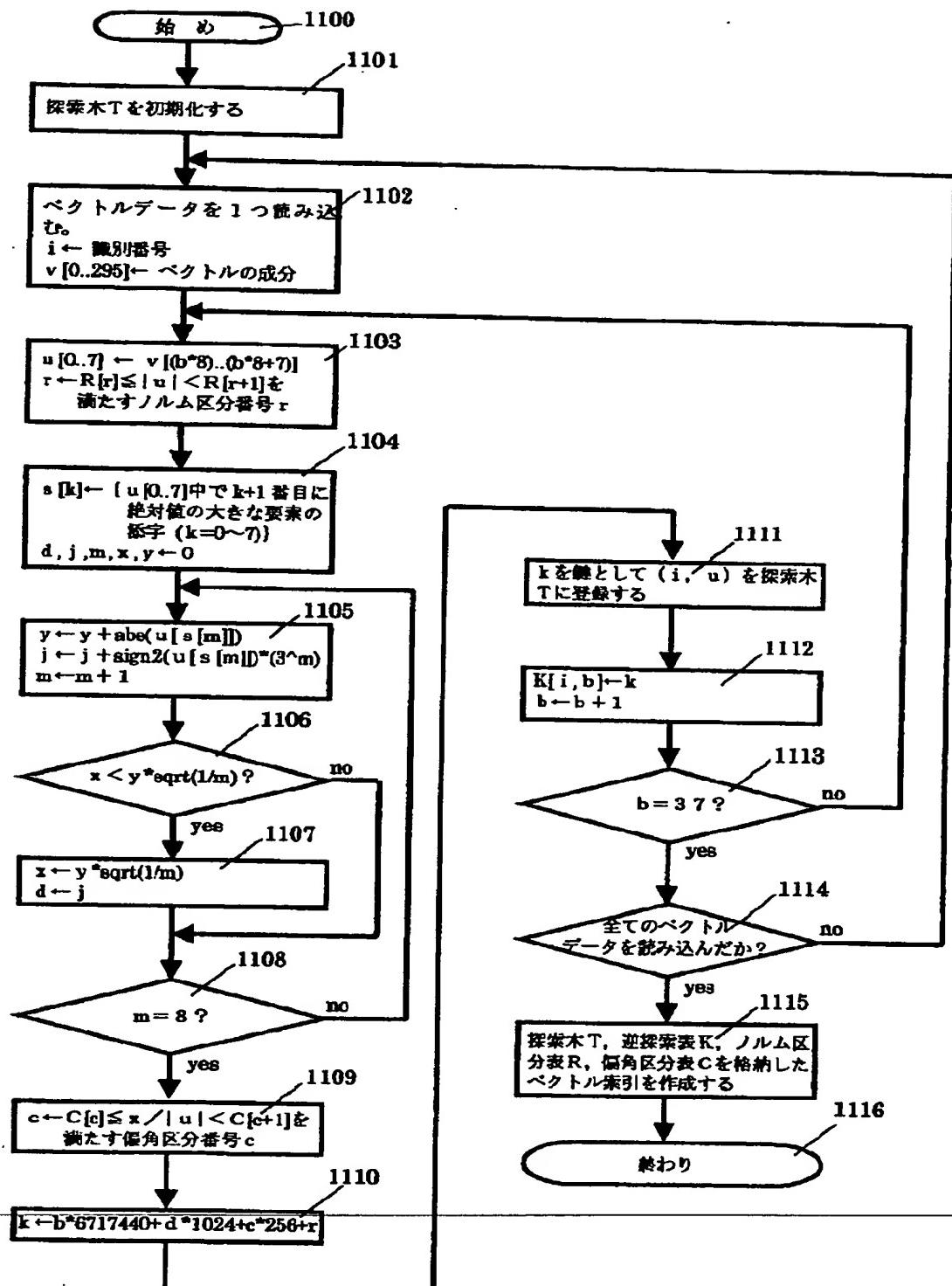
【図4】



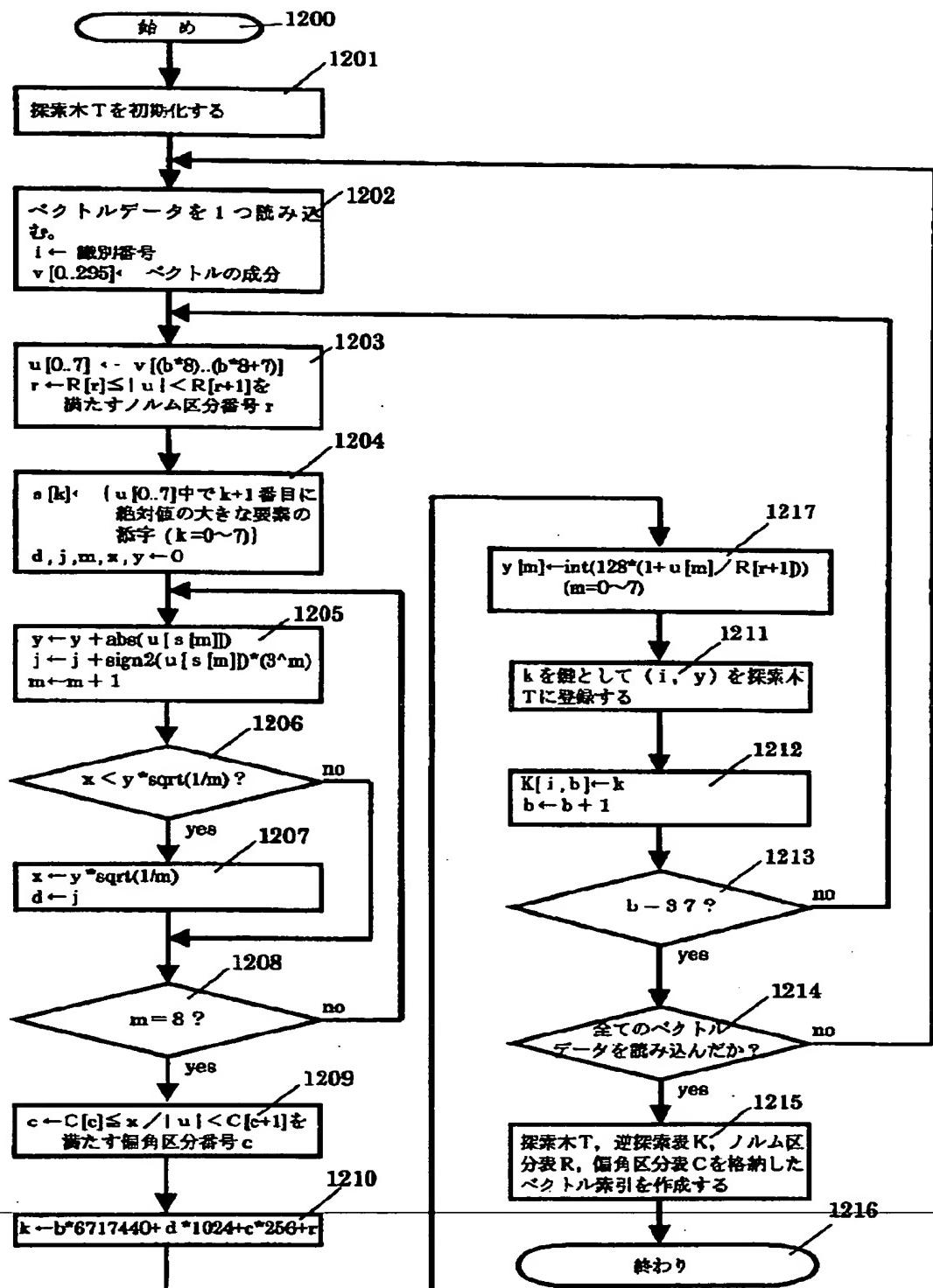
【図5】



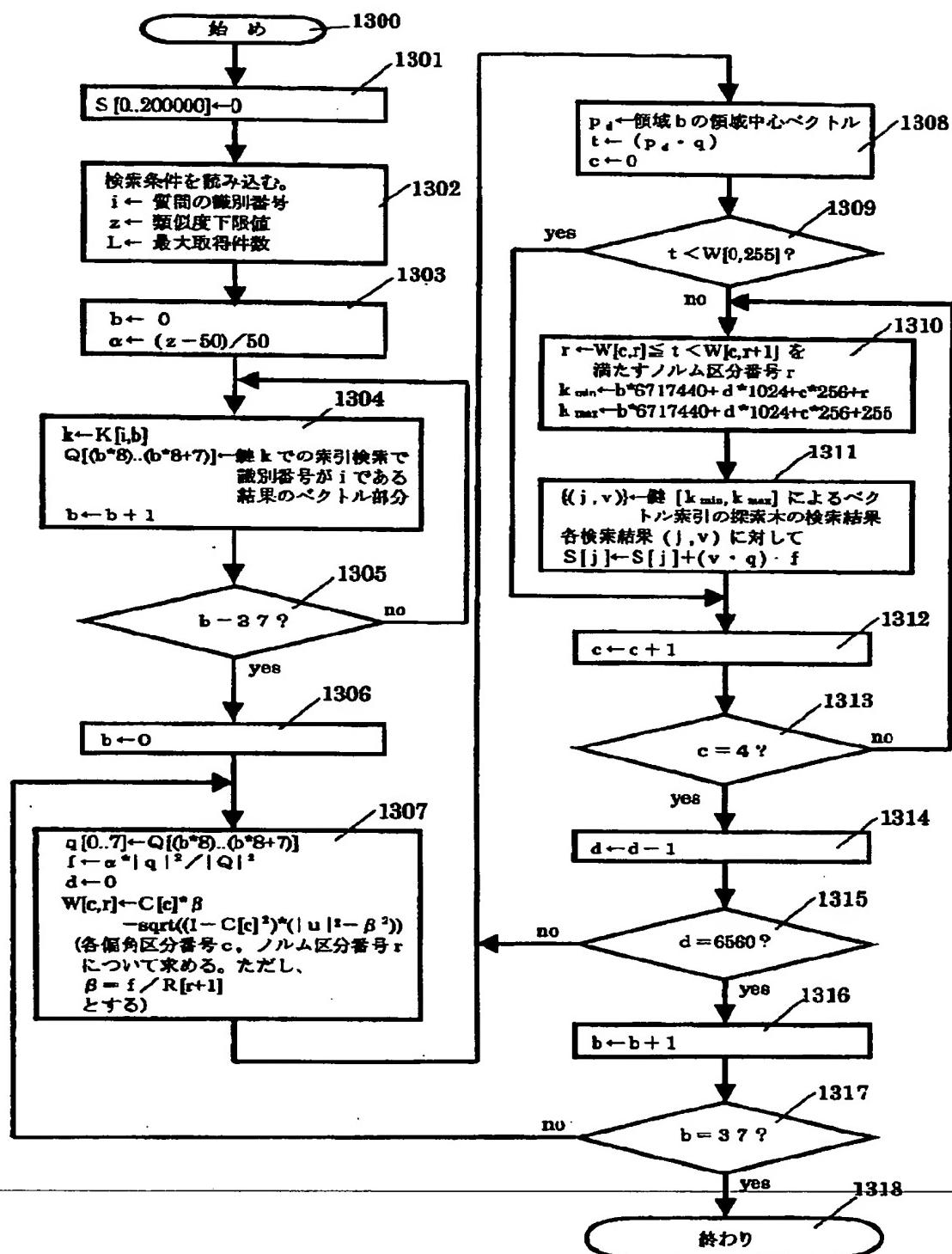
【図6】



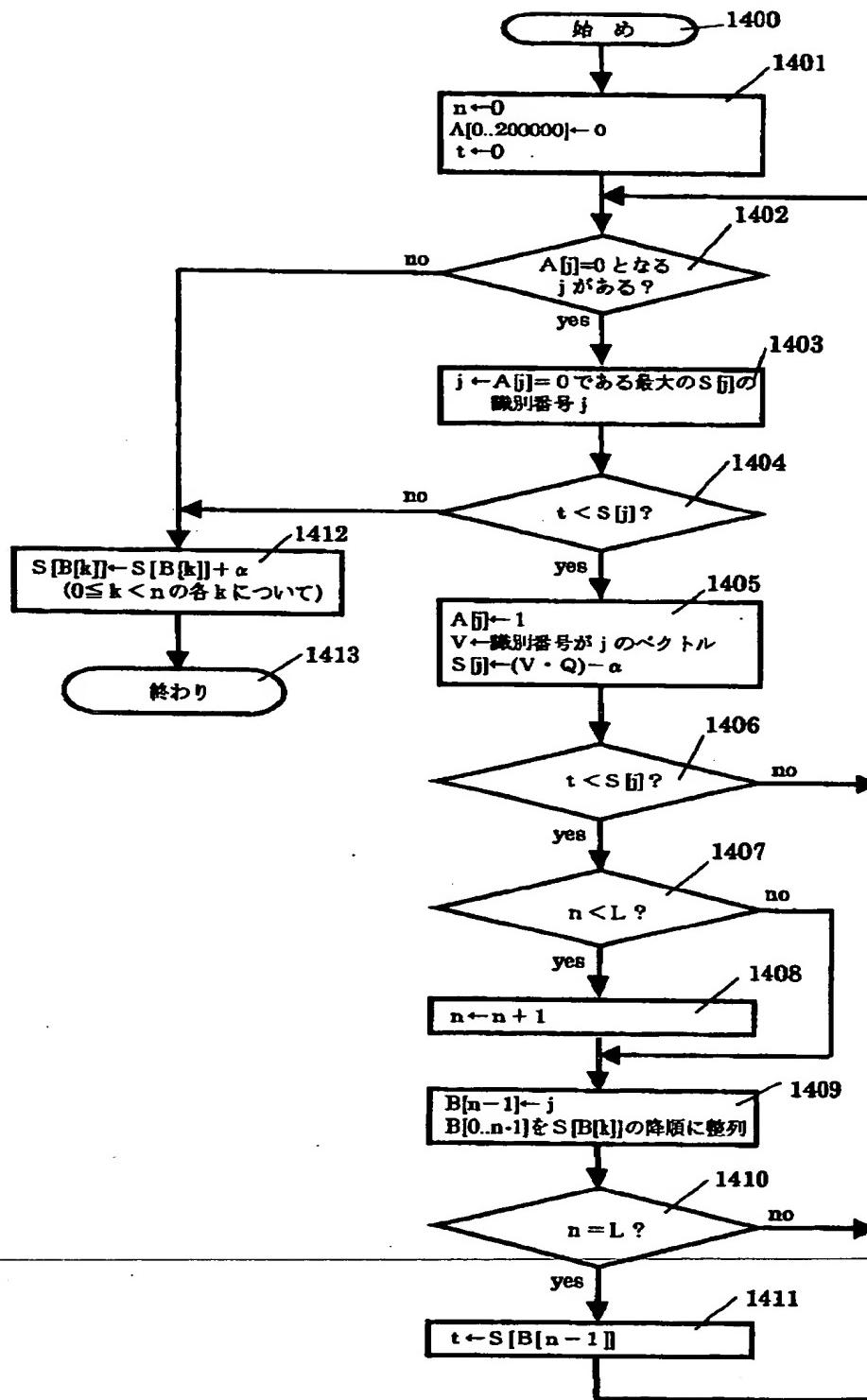
【図7】



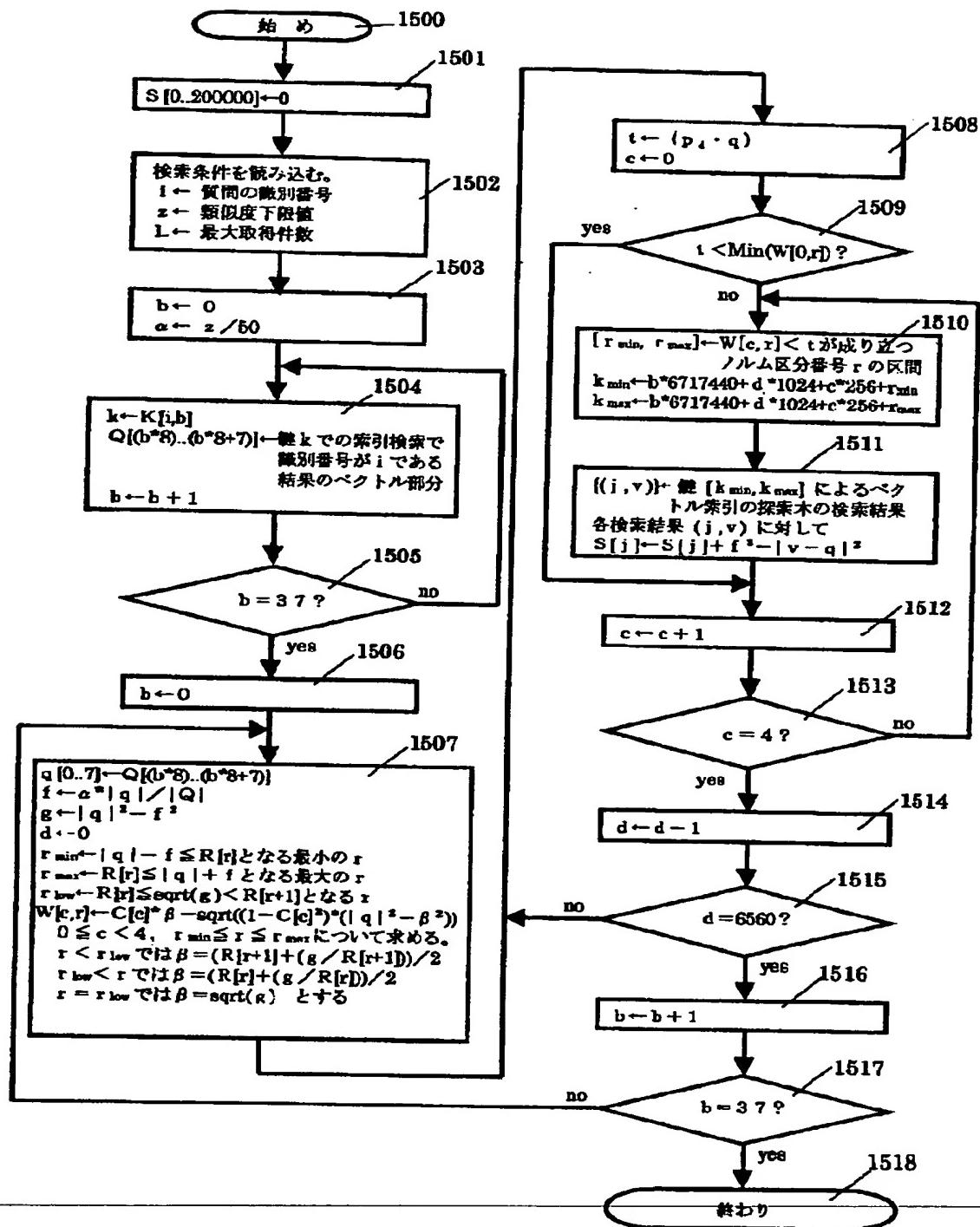
【図8】



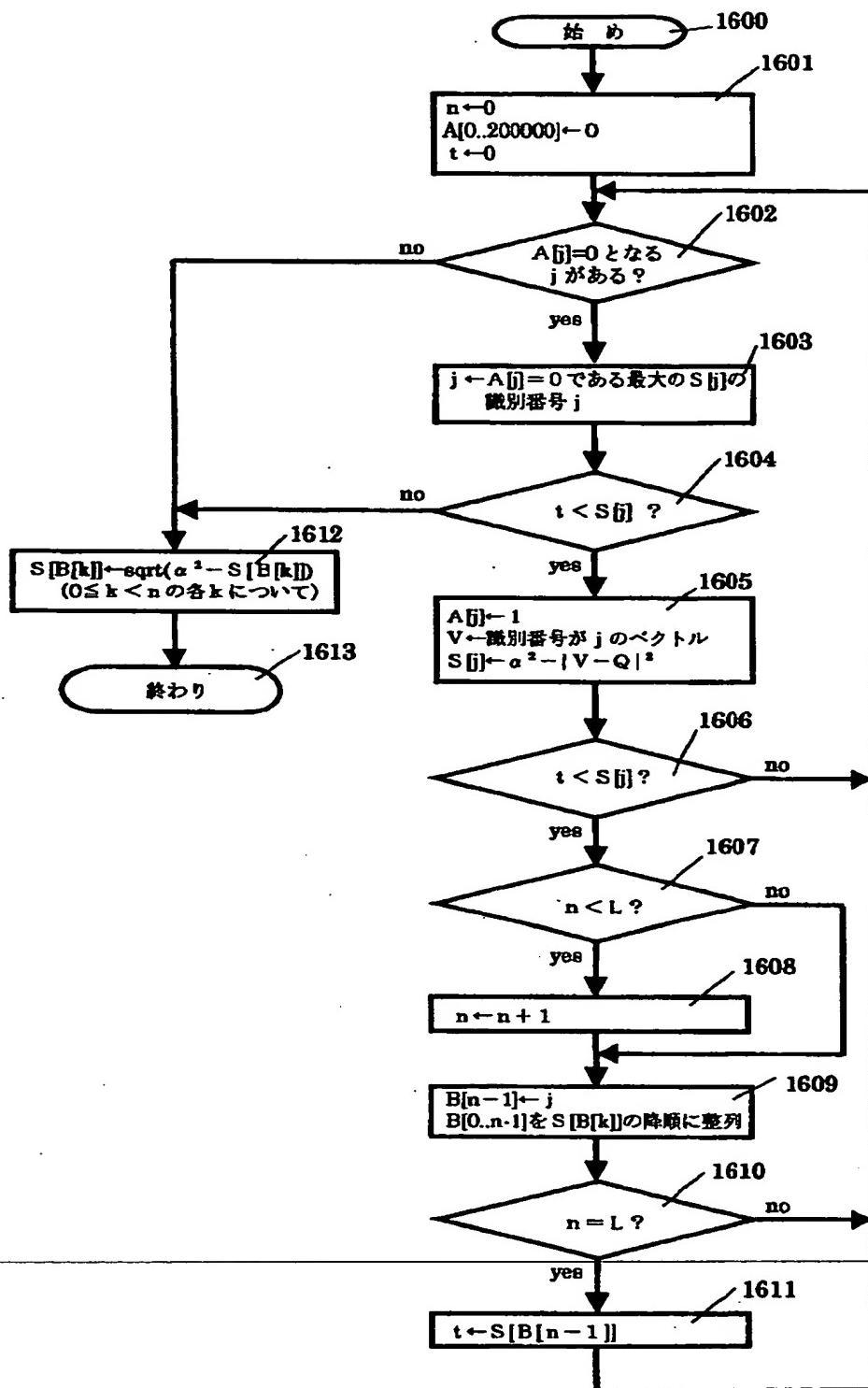
【図9】



【図10】



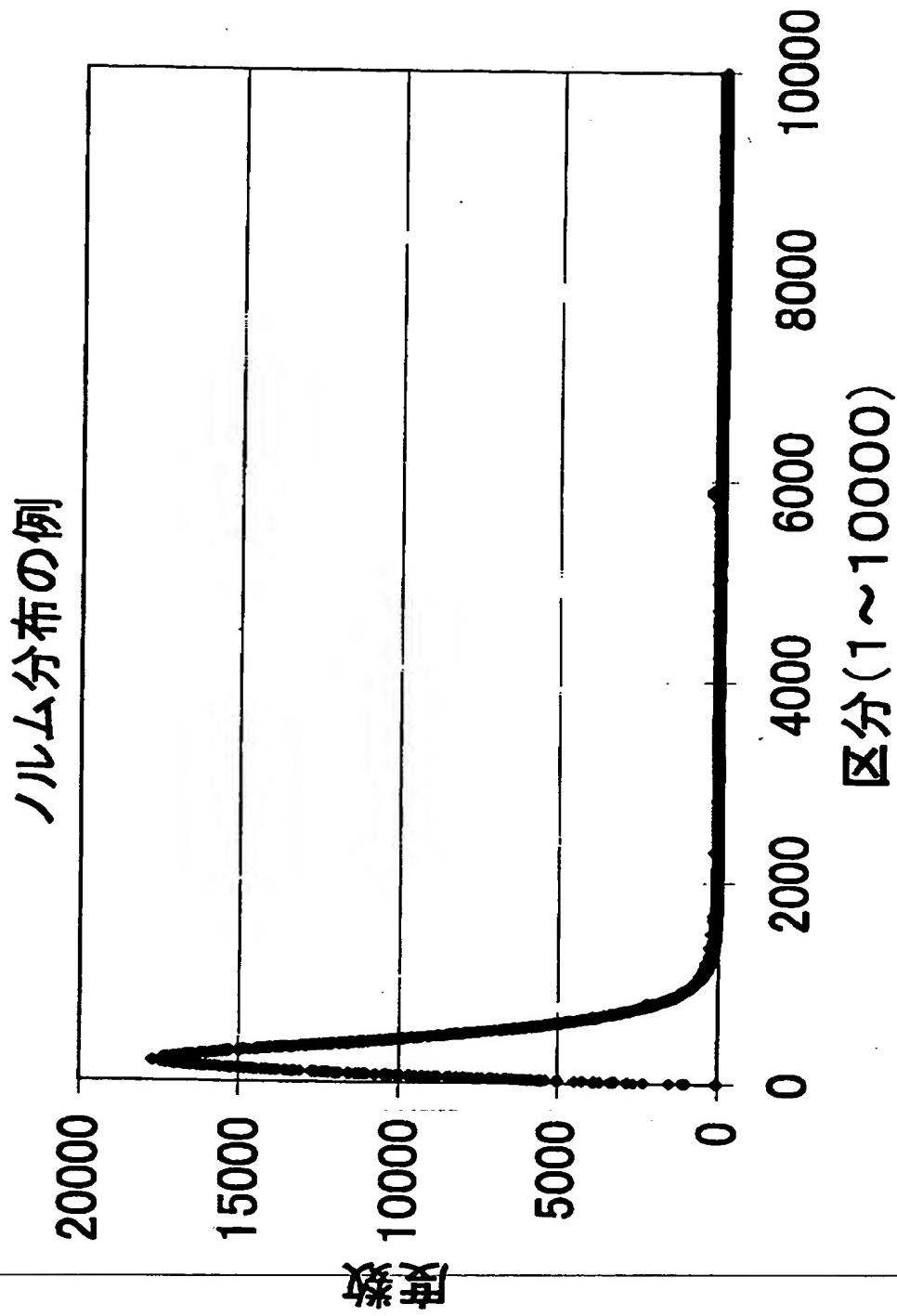
【図11】



【図12】

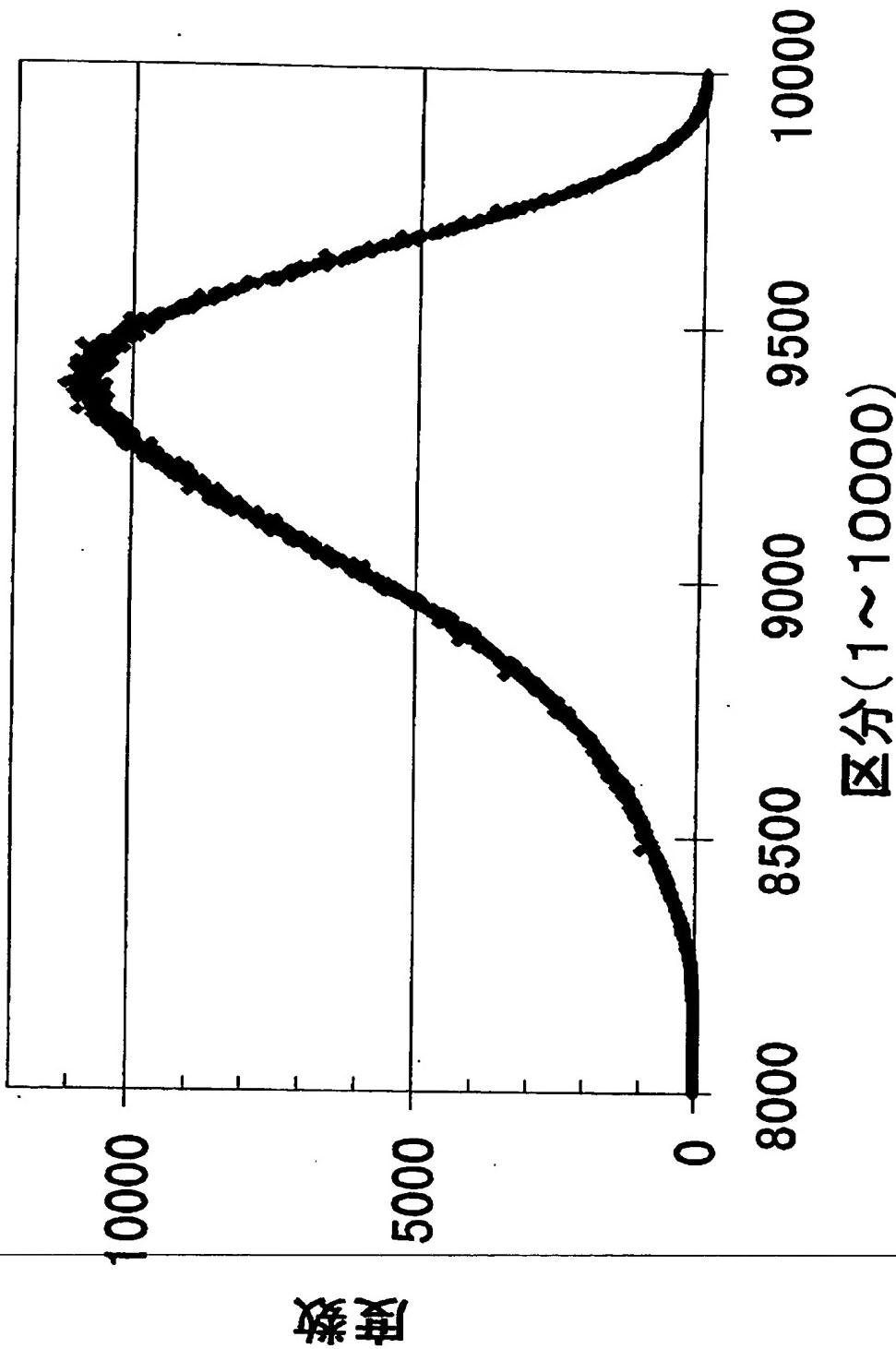
1 +0.029259 -0.016005 -0.021118 +0.024992 -0.006860 -0.009032 -0.007255 -0.007715 -0.025648 +0.016061
 -0.060584 -0.013693 -0.020585 0.112403 -0.012045 +0.044741 +0.026761 +0.078339 +0.048166 +0.043434
 +0.100093 +0.009913 +0.085770 +0.101257 +0.072163 -0.066118 +0.059376 -0.020159 +0.051960 -0.129158
 -0.028065 +0.027535 +0.028316 +0.050490 +0.015931 -0.040316 -0.013109 -0.014728 -0.004639 -0.021525
 -0.000471 -0.033506 +0.013866 -0.054646 +0.067350 +0.042063 +0.041963 -0.006444 -0.092581 +0.004488
 +0.004741 +0.009351 +0.038429 -0.042254 -0.027641 -0.068727 +0.037185 0.003393 -0.040649 +0.013169
 +0.020619 +0.025594 -0.019990 -0.117804 +0.005791 -0.027850 +0.000220 -0.038765 -0.023964 +0.020038
 +0.032435 -0.027518 -0.063942 +0.085381 +0.038776 +0.051395 +0.004047 +0.092011 -0.076222 +0.096729
 0.018331 +0.115754 -0.038478 +0.131147 -0.074560 +0.080634 -0.186932 +0.024004 +0.047046 -0.075571
 +0.121789 -0.055221 -0.001166 -0.053469 -0.086326 +0.011837 -0.060801 +0.222437 -0.055550 -0.117881
 -0.020700 -0.028172 0.121642 -0.160389 +0.147845 -0.037681 -0.057998 +0.104025 +0.251415 -0.029438
 +0.030504 -0.048312 -0.072984 -0.088780 +0.041684 +0.127198 +0.061804 +0.061417 -0.016588 +0.024305
 +0.060558 -0.004070 +0.094040 -0.011500 +0.000545 +0.083231 +0.016565 +0.081034 +0.073438 -0.006857
 -0.008995 +0.023537 +0.066849 -0.035310 +0.005572 -0.015236 +0.109983 0.185597 +0.016643 +0.032632
 -0.075726 -0.110307 +0.038577 +0.038475 -0.042287 +0.082878 +0.035997 -0.009888 +0.081286 +0.063583
 -0.041429 +0.026969 -0.040406 +0.005639 +0.032087 +0.007947 +0.041689 +0.010077 +0.067726 -0.101670
 -0.091163 +0.167914 -0.080320 +0.049351 +0.069409 +0.063139 -0.038358 -0.126212 +0.058109 +0.031847
 0.014998 -0.022995 +0.054876 +0.033124 -0.065283 -0.058574 +0.049729 0.046552 +0.042485 -0.006179
 -0.058764 +0.079383 +0.000817 -0.001482 -0.036410 -0.036097 -0.045920 -0.001729 +0.039971 +0.083165
 -0.023112 +0.014492 +0.028403 +0.047480 +0.038502 -0.028348 +0.055128 -0.045340 -0.066148 +0.018156
 -0.008535 -0.042836 +0.006119 -0.037691 +0.08055 +0.035741 -0.023994 -0.012401 -0.070880 +0.010066
 -0.013264 -0.031192 -0.064061 -0.026757 -0.028246 +0.078634 +0.013295 +0.011129 +0.028807 +0.012339
 +0.007173 -0.006858 +0.040397 +0.039853 +0.085247 -0.063102 +0.052307 0.065223 +0.116747 +0.013464
 -0.004875 +0.019186 +0.024114 -0.056101 -0.024008 +0.061251 -0.043466 -0.017640 +0.081001 -0.014824
 -0.003836 +0.059081 +0.051650 -0.032798 -0.039059 +0.020370 +0.015096 +0.01693 +0.015507 -0.041601
 -0.000192 -0.065087 +0.018487 -0.040416 -0.036173 -0.011809 -0.010862 +0.005944 +0.026534 -0.031335
 +0.023075 +0.033037 +0.063589 +0.014185 +0.006539 +0.002593 -0.023986 -0.038277 -0.009555 -0.018987
 +0.052526 +0.035448 +0.013042 +0.023662 +0.011775 -0.055742 -0.008120 -0.040546 -0.023508 0.069309
 +0.037886 +0.041494 -0.038487 -0.035241 -0.020432 -0.008060 -0.002984 +0.070241 +0.069379 -0.020206
 +0.032996 +0.047815 +0.046106 +0.001794 +0.015342 -0.003895
 2 +0.028972 -0.012757 -0.015597 -0.019727 +0.009386 -0.016593 +0.003627 +0.006288 -0.019184 +0.020305
 -0.057163 -0.017815 -0.026345 -0.102036 +0.002587 +0.037285 +0.029168 +0.076061 +0.043901 +0.040049
 +0.123462 +0.001139 +0.085437 +0.108889 +0.062662 -0.048914 +0.060612 -0.005019 +0.030421 -0.153633
 -0.041444 +0.038908 +0.006823 +0.069954 +0.028216 -0.043207 -0.030092 +0.013758 +0.023770 -0.018313
 +0.008825 +0.036443 +0.001076 -0.067721 +0.046034 +0.030717 +0.017880 -0.036936 -0.093124 -0.000883
 +0.027865 +0.007906 +0.005978 -0.024367 -0.012682 -0.054200 +0.025934 +0.000926 -0.047710 +0.008757
 +0.005940 +0.020855 -0.070890 -0.113381 +0.04988 0.038150 -0.000456 -0.024836 -0.007560 +0.025912
 +0.009004 -0.053047 -0.079142 +0.085440 +0.027876 +0.051104 +0.018944 +0.082277 -0.071359 +0.107303
 +0.005041 +0.112418 -0.009138 +0.119492 -0.069016 +0.123591 -0.168734 +0.032382 +0.005430 -0.030192
 +0.116327 -0.077304 +0.003280 -0.008984 -0.055858 +0.022018 -0.110375 +0.197565 -0.038060 -0.065170
 -0.065823 -0.021350 -0.104387 -0.147696 +0.111377 -0.028878 -0.097095 +0.064212 +0.255376 -0.011009
 +0.020901 -0.032671 -0.092766 -0.063843 +0.008917 +0.106446 +0.070094 +0.078741 -0.012886 -0.003581
 +0.069363 +0.021164 -0.046900 -0.021002 -0.008879 +0.052981 +0.006370 +0.081378 +0.054328 -0.006424
 -0.006277 +0.013635 -0.117156 -0.037470 +0.014036 -0.048765 +0.093100 0.147319 +0.028556 -0.017833
 -0.070005 -0.123845 +0.013978 +0.006964 -0.047420 +0.100905 -0.019278 -0.009641 +0.057287 +0.058665
 -0.063796 +0.019097 -0.045014 -0.036129 +0.022014 +0.071405 +0.028573 +0.046653 +0.063911 -0.048555
 -0.070203 +0.205558 0.051782 +0.102727 +0.042066 +0.028359 -0.021939 -0.082837 +0.064817 +0.017215
 -0.042670 -0.031901 +0.037476 +0.055012 -0.012237 -0.067371 +0.072587 0.009949 +0.053991 +0.019722
 -0.035742 +0.081726 +0.019732 +0.013624 -0.031871 -0.009025 -0.061237 +0.002162 +0.014326 +0.103617
 -0.048376 +0.034422 -0.019797 +0.041018 +0.089878 +0.072000 +0.030657 +0.040709 -0.071603 +0.005629
 .
 .

【図13】



【図14】

偏角分布の例



【図15】

	0	1	2	3	4	5	6	7	8	9
0	0.00000	0.03320	0.04112	0.04617	0.04999	0.05354	0.05687	0.05940	0.06182	0.06416
10	0.06585	0.06805	0.06965	0.07121	0.07275	0.07425	0.07572	0.07716	0.07858	0.07951
20	0.08088	0.08178	0.08312	0.08400	0.08530	0.08616	0.08701	0.08827	0.08910	0.08992
30	0.09073	0.09194	0.09273	0.09352	0.09431	0.09508	0.09585	0.09662	0.09737	0.09813
40	0.09887	0.09961	0.10035	0.10108	0.10180	0.10252	0.10324	0.10395	0.10465	0.10535
50	0.10605	0.10639	0.10708	0.10776	0.10844	0.10912	0.10979	0.11046	0.11079	0.11145
60	0.11211	0.11276	0.11309	0.11374	0.11438	0.11502	0.11566	0.11598	0.11661	0.11724
70	0.11786	0.11849	0.11879	0.11941	0.12003	0.12064	0.12094	0.12155	0.12215	0.12275
80	0.12305	0.12365	0.12424	0.12483	0.12542	0.12571	0.12629	0.12687	0.12745	0.12774
90	0.12831	0.12888	0.12945	0.13002	0.13030	0.13087	0.13143	0.13198	0.13254	0.13282
100	0.13337	0.13392	0.13447	0.13501	0.13556	0.13583	0.13637	0.13690	0.13744	0.13797
110	0.13851	0.13904	0.13956	0.13983	0.14035	0.14087	0.14139	0.14191	0.14243	0.14295
120	0.14346	0.14397	0.14448	0.14499	0.14549	0.14600	0.14650	0.14700	0.14750	0.14800
130	0.14849	0.14899	0.14948	0.14997	0.15046	0.15095	0.15143	0.15192	0.15240	0.15288
140	0.15336	0.15408	0.15456	0.15503	0.15551	0.15598	0.15645	0.15715	0.15762	0.15808
150	0.15878	0.15924	0.15970	0.16016	0.16085	0.16131	0.16199	0.16244	0.16289	0.16357
160	0.16402	0.16469	0.16513	0.16580	0.16624	0.16690	0.16734	0.16800	0.16866	0.16909
170	0.16974	0.17039	0.17104	0.17147	0.17211	0.17275	0.17338	0.17402	0.17465	0.17507
180	0.17570	0.17633	0.17716	0.17778	0.17840	0.17902	0.17963	0.18024	0.18106	0.18166
190	0.18227	0.18308	0.18368	0.18447	0.18507	0.18586	0.18665	0.18724	0.18803	0.18881
200	0.18958	0.19036	0.19113	0.19190	0.19266	0.19342	0.19437	0.19512	0.19606	0.19681
210	0.19774	0.19867	0.19959	0.20051	0.20143	0.20252	0.20342	0.20450	0.20540	0.20647
220	0.20754	0.20860	0.20983	0.21087	0.21209	0.21330	0.21450	0.21587	0.21706	0.21858
230	0.21992	0.22142	0.22291	0.22438	0.22602	0.22780	0.22957	0.23148	0.23338	0.23557
240	0.23774	0.24005	0.24249	0.24520	0.24818	0.25142	0.25505	0.25919	0.26369	0.26921
250	0.27595	0.28434	0.29600	0.31512	0.35936	0.49100	0.85733			

【図16】

偏角区分表の例 (4区分)	
区分番号	区分境界
0	0.8274
1	0.9079
2	0.9301
3	0.9486
4	1.0000

【図17】

区分	0	1	2	3	4	5	6	7	8	9
b = 0, c = 0										
0	+9.99999	+0.03142	+0.01968	+0.01436	+0.01008	+0.00750	+0.00536	+0.00355	+0.00220	+0.00100
10	+0.00010	-0.00089	-0.00164	-0.00233	-0.00298	-0.00347	-0.00404	-0.00447	-0.00498	-0.00537
20	-0.00574	-0.00609	-0.00643	-0.00675	-0.00706	-0.00736	-0.00764	-0.00792	-0.00812	-0.00838
30	-0.00857	-0.00881	-0.00904	-0.00921	-0.00938	-0.00959	-0.00975	-0.00996	-0.01010	-0.01025
40	-0.01039	-0.01058	-0.01071	-0.01084	-0.01097	-0.01110	-0.01122	-0.01139	-0.01153	-0.01162
50	-0.01173	-0.01185	-0.01196	-0.01206	-0.01217	-0.01227	-0.01237	-0.01247	-0.01257	-0.01267
60	-0.01276	-0.01286	-0.01292	-0.01301	-0.01310	-0.01319	-0.01327	-0.01336	-0.01344	-0.01352
70	-0.01360	-0.01368	-0.01376	-0.01381	-0.01389	-0.01397	-0.01404	-0.01411	-0.01418	-0.01426
80	-0.01433	-0.01439	-0.01446	-0.01453	-0.01457	-0.01464	-0.01470	-0.01477	-0.01483	-0.01489
90	-0.01496	-0.01502	-0.01508	-0.01514	-0.01519	-0.01525	-0.01531	-0.01536	-0.01542	-0.01547
100	-0.01553	-0.01558	-0.01564	-0.01569	-0.01574	-0.01579	-0.01584	-0.01589	-0.01594	-0.01599
110	-0.01604	-0.01609	-0.01613	-0.01618	-0.01623	-0.01627	-0.01633	-0.01638	-0.01642	-0.01646
120	-0.01651	-0.01656	-0.01661	-0.01665	-0.01669	-0.01673	-0.01678	-0.01682	-0.01686	-0.01692
130	-0.01695	-0.01699	-0.01704	-0.01708	-0.01712	-0.01717	-0.01723	-0.01725	-0.01729	-0.01733
140	-0.01737	-0.01741	-0.01745	-0.01749	-0.01753	-0.01757	-0.01761	-0.01766	-0.01769	-0.01773
150	-0.01777	-0.01780	-0.01784	-0.01788	-0.01792	-0.01796	-0.01803	-0.01804	-0.01808	-0.01812
160	-0.01816	-0.01819	-0.01823	-0.01827	-0.01830	-0.01835	-0.01838	-0.01842	-0.01846	-0.01849
170	-0.01854	-0.01857	-0.01861	-0.01864	-0.01868	-0.01872	-0.01875	-0.01879	-0.01883	-0.01887
180	-0.01891	-0.01894	-0.01898	-0.01902	-0.01906	-0.01909	-0.01913	-0.01917	-0.01920	-0.01924
190	-0.01928	-0.01932	-0.01936	-0.01939	-0.01943	-0.01947	-0.01951	-0.01955	-0.01958	-0.01962
200	-0.01966	-0.01970	-0.01974	-0.01978	-0.01982	-0.01986	-0.01990	-0.01994	-0.01998	-0.02002
210	-0.02006	-0.02010	-0.02015	-0.02019	-0.02023	-0.02027	-0.02032	-0.02036	-0.02040	-0.02045
220	-0.02049	-0.02054	-0.02059	-0.02063	-0.02068	-0.02073	-0.02078	-0.02083	-0.02086	-0.02094
230	-0.02099	-0.02104	-0.02110	-0.02116	-0.02121	-0.02128	-0.02134	-0.02140	-0.02147	-0.02154
240	-0.02161	-0.02169	-0.02177	-0.02185	-0.02194	-0.02204	-0.02215	-0.02227	-0.02241	-0.02256
250	-0.02275	-0.02299	-0.02334	-0.02401	-0.02527					
b = 0, c = 3										
0	+9.99999	+0.04126	+0.03220	+0.02771	+0.02395	+0.02162	+0.01966	+0.01793	+0.01672	+0.01559
10	+0.01473	+0.01378	+0.01306	+0.01239	+0.01176	+0.01129	+0.01073	+0.01030	+0.00980	+0.00942
20	+0.00905	+0.00871	+0.00837	+0.00805	+0.00774	+0.00744	+0.00715	+0.00688	+0.00668	+0.00642
30	+0.00623	+0.00598	+0.00575	+0.00558	+0.00541	+0.00519	+0.00503	+0.00482	+0.00467	+0.00452
40	+0.00438	+0.00419	+0.00405	+0.00392	+0.00378	+0.00365	+0.00353	+0.00336	+0.00324	+0.00312
50	+0.00300	+0.00289	+0.00278	+0.00267	+0.00256	+0.00245	+0.00235	+0.00224	+0.00214	+0.00204
60	+0.00194	+0.00185	+0.00178	+0.00169	+0.00160	+0.00151	+0.00142	+0.00133	+0.00124	+0.00116
70	+0.00107	+0.00099	+0.00091	+0.00085	+0.00077	+0.00070	+0.00062	+0.00054	+0.00047	+0.00039
80	+0.00032	+0.00025	+0.00018	+0.00010	+0.00006	-0.00001	-0.00008	-0.00015	-0.00021	-0.00028
90	-0.00034	-0.00040	-0.00047	-0.00053	-0.00059	-0.00065	-0.00071	-0.00077	-0.00083	-0.00089
100	-0.00094	-0.00100	-0.00106	-0.00111	-0.00117	-0.00122	-0.00127	-0.00133	-0.00138	-0.00143
110	-0.00148	-0.00153	-0.00158	-0.00163	-0.00168	-0.00173	-0.00179	-0.00184	-0.00189	-0.00193
120	-0.00198	-0.00204	-0.00208	-0.00213	-0.00217	-0.00221	-0.00227	-0.00231	-0.00236	-0.00241
130	-0.00245	-0.00249	-0.00255	-0.00259	-0.00263	-0.00268	-0.00272	-0.00277	-0.00281	-0.00286
140	-0.00289	-0.00294	-0.00298	-0.00303	-0.00306	-0.00311	-0.00316	-0.00320	-0.00324	-0.00328
150	-0.00333	-0.00336	-0.00340	-0.00345	-0.00349	-0.00353	-0.00357	-0.00361	-0.00366	-0.00370
160	-0.00374	-0.00378	-0.00382	-0.00386	-0.00389	-0.00394	-0.00398	-0.00402	-0.00406	-0.00410
170	-0.00414	-0.00418	-0.00423	-0.00426	-0.00430	-0.00434	-0.00438	-0.00442	-0.00446	-0.00451
180	-0.00455	-0.00458	-0.00463	-0.00467	-0.00470	-0.00474	-0.00478	-0.00483	-0.00486	-0.00491
190	-0.00494	-0.00499	-0.00503	-0.00507	-0.00511	-0.00515	-0.00519	-0.00523	-0.00527	-0.00532
200	-0.00536	-0.00540	-0.00544	-0.00548	-0.00553	-0.00557	-0.00562	-0.00566	-0.00571	-0.00575
210	-0.00579	-0.00584	-0.00589	-0.00593	-0.00598	-0.00602	-0.00607	-0.00612	-0.00617	-0.00622
220	-0.00627	-0.00632	-0.00637	-0.00642	-0.00647	-0.00653	-0.00658	-0.00663	-0.00669	-0.00675
230	-0.00680	-0.00686	-0.00692	-0.00699	-0.00705	-0.00712	-0.00719	-0.00726	-0.00733	-0.00741
240	-0.00749	-0.00757	-0.00766	-0.00775	-0.00786	-0.00797	-0.00808	-0.00821	-0.00837	-0.00854
250	-0.00875	-0.00901	-0.01015	-0.01157						

【図18】

区分	0	1	2	3	4	5	6	7	8	9
b=0, c=0										
0 +9.99999 +0.03665 +0.03356 +0.03363 +0.03460 +0.03613 +9.99999 +9.99999 +9.99999 +9.99999										
b=0, c=1										
0 +9.99999 +0.03869 +0.03581 +0.03588 +0.03678 +0.03821 +9.99999 +9.99999 +9.99999 +9.99999										
b=0, c=2										
0 +9.99999 +0.04164 +0.03912 +0.03916 +0.03997 +0.04122 +9.99999 +9.99999 +9.99999 +9.99999										
b=0, c=3										
0 +9.99999 +0.04476 +0.04275 +0.04280 +0.04344 +0.04444 +9.99999 +9.99999 +9.99999 -9.99999										
b=1, c=0										
60 +9.99999 +9.99999 +9.99999 +0.11286 +0.11079 +0.10939 +0.10827 +0.10731 +0.10648 +0.10573										
70 +0.10506 +0.10444 +0.10387 +0.10335 +0.10302 +0.10255 +0.10212 +0.10171 +0.10133 +0.10097										
80 +0.10064 +0.10032 +0.10003 +0.09975 +0.09949 +0.09932 +0.09909 +0.09887 +0.09866 +0.09846										
90 +0.09828 +0.09811 +0.09795 +0.09781 +0.09767 +0.09754 +0.09742 +0.09731 +0.09722 +0.09712										
100 -0.09704 +0.09697 +0.09690 +0.09685 +0.09679 +0.09675 +0.09671 +0.09669 +0.09666 +0.09665										
110 +0.09663 +0.09663 +0.09664 +0.09665 +0.09666 +0.09668 +0.09671 +0.09675 +0.09679 +0.09684										
120 +0.09689 +0.09694 +0.09703 +0.09709 +0.09717 +0.09725 +0.09733 +0.09745 +0.09754 +0.09764										
130 +0.09778 +0.09789 +0.09800 +0.09816 +0.09829 +0.09842 +0.09860 +0.09875 +0.09894 +0.09910										
140 +0.09931 +0.09947 +0.09970 +0.09988 +0.10012 +0.10031 +0.10057 +0.10084 +0.10112 +0.10134										
150 +0.10163 +0.10194 +0.10218 +0.10252 +0.10286 +0.10322 +0.10359 +0.10397 +0.10438 +0.10480										
160 +0.10524 +0.10571 +0.10620 +0.10672 +0.10727 +0.10787 +0.10868 +0.10941 +0.11023 +0.11149										
170 +0.11287 +9.99999 +9.99999 +9.99999 +9.99999 +9.99999 +9.99999 +9.99999 +9.99999 +9.99999										
b=1, c=1										
60 +9.99999 +9.99999 +9.99999 +0.11816 +0.11626 +0.11497 +0.11398 +0.11306 +0.11229 +0.11160										
70 +0.11098 +0.11040 +0.10987 +0.10939 +0.10908 +0.10865 +0.10824 +0.10786 +0.10751 +0.10718										
80 +0.10686 +0.10657 +0.10629 +0.10604 +0.10579 +0.10564 +0.10542 +0.10521 +0.10502 +0.10483										
90 +0.10466 +0.10450 +0.10436 +0.10422 +0.10409 +0.10397 +0.10386 +0.10376 +0.10366 +0.10358										
100 +0.10350 +0.10343 +0.10337 +0.10332 +0.10327 +0.10323 +0.10320 +0.10317 +0.10315 +0.10313										
110 +0.10312 +0.10312 +0.10313 +0.10313 +0.10317 +0.10319 +0.10323 +0.10327 +0.10331										
120 +0.10336 +0.10341 +0.10349 +0.10355 +0.10362 +0.10369 +0.10377 +0.10388 +0.10397 +0.10406										
130 +0.10419 +0.10430 +0.10440 +0.10455 +0.10467 +0.10479 +0.10496 +0.10510 +0.10528 +0.10543										
140 +0.10562 +0.10578 +0.10599 +0.10616 +0.10638 +0.10656 +0.10680 +0.10705 +0.10731 +0.10752										
150 +0.10779 +0.10808 +0.10831 +0.10861 +0.10893 +0.10927 +0.10961 +0.10997 +0.11034 +0.11074										
160 +0.11115 +0.11158 +0.11203 +0.11251 +0.11302 +0.11357 +0.11432 +0.11499 +0.11575 +0.11690										
170 +0.11817 +9.99999 +9.99999 +9.99999 +9.99999 +9.99999 +9.99999 +9.99999 +9.99999 +9.99999										
b=1, c=2										
60 +9.99999 +9.99999 +9.99999 +0.12564 +0.12403 +0.12294 +0.12206 +0.12131 +0.12065 +0.12006										
70 +0.11951 +0.11902 +0.11856 +0.11814 +0.11787 +0.11749 +0.11714 +0.11681 +0.11650 +0.11621										
80 +0.11594 +0.11568 +0.11544 +0.11521 +0.11500 +0.11486 +0.11467 +0.11449 +0.11431 +0.11415										
90 +0.11400 +0.11386 +0.11373 +0.11361 +0.11350 +0.11339 +0.11329 +0.11321 +0.11312 +0.11305										
100 +0.11298 +0.11292 +0.11286 +0.11282 +0.11277 +0.11274 +0.11271 +0.11268 +0.11267 +0.11265										
110 +0.11264 +0.11264 +0.11265 +0.11265 +0.11266 +0.11266 +0.11270 +0.11274 +0.11277 +0.11281										
120 +0.11285 +0.11290 +0.11297 +0.11302 +0.11308 +0.11315 +0.11322 +0.11331 +0.11339 +0.11347										
130 +0.11359 +0.11368 +0.11377 +0.11391 +0.11401 +0.11412 +0.11427 +0.11439 +0.11455 +0.11467										
140 +0.11485 +0.11498 +0.11517 +0.11532 +0.11552 +0.11567 +0.11588 +0.11610 +0.11633 +0.11651										
150 +0.11675 +0.11700 +0.11720 +0.11747 +0.11774 +0.11803 +0.11833 +0.11864 +0.11897 +0.11931										
160 +0.11966 +0.12003 +0.12043 +0.12084 +0.12128 +0.12175 +0.12239 +0.12296 +0.12361 +0.12458										
170 +0.12564 +9.99999 +9.99999 +9.99999 +9.99999 +9.99999 +9.99999 +9.99999 +9.99999 +9.99999										

【書類名】 要約書

【要約】

【課題】 数百次元のベクトルデータベースから、単一のベクトル索引によつて、内積あるいは距離のいずれかの尺度による、類似検索範囲と最大取得件数を指定した類似ベクトル検索を高速に行う。

【解決手段】 ベクトル索引作成の際には、各ベクトルを複数の部分ベクトルに分解し、ノルム区分と所属領域と偏角区分とによって特徴付けることで索引を作成し、類似検索の際には、質問のベクトルと検索範囲とから部分質問ベクトルと部分検索範囲を求めて、各部分空間で類似検索を行って検索範囲との差分を累算して上限値を求め、上限値の高いものから正確な尺度を求めていって最終的な類似検索結果を得る。

【選択図】 図1

出願人履歴情報

識別番号 [000005821]

1. 変更年月日 1990年 8月28日

[変更理由] 新規登録

住 所 大阪府門真市大字門真1006番地

氏 名 松下電器産業株式会社

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)